Journal of the

AMERICAN STATISTICAL ASSOCIATION

JUN

JUNE 1954

Measurement for Economic Models ..... Stanley Lebergott

Technical Aspects of Transportation Flow Data ..... A. Tyner Smith

Response Errors in the Collection of Wage Statistics by Mail Questionnaires ..... Samuel E. Cohen and Benjamin Lev

Industrial Classes in the United States, 1870 to 1950 ..... Philip M. Hauser

Statistical Methods for Poisson Processes and Exponential Populations ..... Z. W. Birnbaum

Applications of the Circular Normal Distribution ..... A. J. Stam

A New Type of Control Chart Limits for Means, Ranges, and Standard Deviations ..... E. S. Pearson

Proceedings of the 113th Annual Meeting .....

Summaries of Papers Delivered at the 113th Annual Meeting .....

Statistical Abstracts .....

BOOK REVIEWS by Kenneth J. Arrow, Geoffrey Beall, Dorothy S. Brady, J. W. Curtiss, John Folger, Sophia Gogek, H. C. Hamaker, Palmer O. Johnson, Oscar Kempthorne, Sherman J. Maisel, Harold Nisselson, H. W. Norton, L. J. Norton, Henry S. Shryock, Jr., Z. Szatrowski, Leona E. Tyler, John E. Walsh, and Bernard L. Welch .....

Random Digits .....

VOLUME 49 NUMBER

American Statistical Association

Organized November 27, 1839, Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and others who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

Annual Dues*

Residents of North America	$ 8	Member's subscription to Biometrics	4
Others	5	Contributing membership	25
Student membership	4	Institutional membership, minimum	100
Introductory membership (first dues payment of members under 30 years old)	6		

* Includes subscription to the *Journal of the American Statistical Association*, and to *The American Statistician*.

Subscription rate, $5 per year. Prices for back issues available on request. A cumulative Index to Volumes 1-34, 1839-1939, may be obtained from the Secretary.

For further information about the Association and membership application forms, write the Secretary, 1108 16th Street, N. W., Washington 6, D.C.

The present Institutional Members are:

American Telephone & Telegraph Company
The Atlantic Refining Company
Armour & Company
Armstrong Cork Company
Chrysler Corporation
The Columbia Gas System, Inc.
Deere & Company
The Dow Chemical Company
Dun & Bradstreet, Inc.
The Equitable Life Assurance Society of the U. S.
Ford Motor Company
General Electric Company
General Foods Corporation
General Motors Corporation
Industrial Surveys Company
International Business Machines
Kennecott Copper Corporation
Metropolitan Life Insurance Company
National Analysts, Inc.
National Bank of Detroit
National Industrial Conference Board
Northwest Bancorporation
Philadelphia Electric Company
Chas. Pfizer & Co.
Public Service Electric & Gas Company
Sears Roebuck & Company
Socony-Vacuum Oil Company, Inc.
Standard Oil Company (Indiana)
Standard Oil Company of New Jersey
Sun Oil Company
United States Steel Corporation
Western Electric Company, Inc.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: 207 Social Hall, University of Chicago, Chicago 37, Illinois. Business Office: 1108 16th St., N.W., Washington 6, D.C. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 412, P. L. & R., authorized March 19, 1928. Entered as second class matter at the post office at Menasha, Wisconsin.

Anyone wishing to change his address should allow eight weeks for it. A copy of the address taken from an issue of this publication should accompany the request of notification.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 49 JUNE, 1954 NUMBER 266

ARTICLES

Measurement for Economic Models . . . . . STANLEY LEBERGOTT 209
Technical Aspects of Transportation Flow Data . . R. TYNES SMITH 227
Response Errors in the Collection of Wage Statistics by Mail Questionnaire . . . . . SAMUEL E. COHEN AND BENJAMIN LIPSTEIN 240
Industrial Classes in the United States, 1870 to 1950 . . . . . TILLMAN M. SOGGE 251
Statistical Methods for Poisson Processes and Exponential Populations. . . . . ALLAN BIRNBAUM 254
Applications of the Circular Normal Distribution . . E. J. GUMBEL 267
A New Type of Control Chart Limits for Means, Ranges, and Sequential Runs . . . . . H. WEILER 298
Proceedings of the 113th Annual Meeting 315
Summaries of Papers Delivered at the 113th Annual Meeting 323
Statistical Abstracts 371

BOOK REVIEWS

GOEDICKE, VICTOR, *Introduction to the Theory of Statistics*. . . . . BERNARD L. WELCH 378
WALKER, HELEN M., AND LEV, JOSEPH, *Statistical Inference* . . . . . PALMER O. JOHNSON 379
LACEY, OLIVER L., *Statistical Methods in Experimentation: An Introduction* . . . . . OSCAR KEMPTHORNE 383
LEWIS, EDWARD E., *Methods of Statistical Analysis in Economics and Business* . . . . . Z. SZATROWSKI 385
BOWKER, A. H., AND GOODE, H. P., *Sampling Inspection by Variables* . . . . . H. C. HAMAKER 386
FINNEY, D. J., *An Introduction to Statistical Science in Agriculture* . . . . . H. W. NORTON 389
THOMSEN, FREDERICK LUNDY, AND FOOTE, RICHARD JAY, *Agricultural Prices* . . . . . L. J. NORTON 391
KLEIN, LAWRENCE R., *A Textbook of Econometrics* . KENNETH J. ARROW 393
DEANE, PHYLLIS, Ed., *Bibliography on Income and Wealth, Volume II, 1948–1949* . . . . . DOROTHY S. BRADY 395
GREBLER, LEO, *The Role of Federal Aids in Residential Construction* . . . . . SHERMAN J. MAISEL 399
MARYLAND, UNIVERSITY OF, *Retail Prices and the Consumer Preference* . . . . . SOPHIA GOGEK 400
MACNIECE, E. H., *Industrial Specifications* . . . J. H. CURTISS 400
CHILTON, NEAL W., *Analysis in Dental Research* . GEOFFREY BEALL 401
HAWLEY, AMOS H., *Intrastate Migration in Michigan: 1935–1940* . . . . . JOHN FOLGER 402
UNITED NATIONS, STATISTICAL OFFICE OF, *Sample Surveys of Current Interest* . . . . . HAROLD NISSELSON 402
LEHMAN, HARVEY C., *Age and Achievement* . . . LEONA E. TYLER 404
HATT, PAUL K., *Backgrounds of Human Fertility in Puerto Rico: A Sociological Survey*. . . . . HENRY S. SHRYOCK, JR. 406
FRENCH, DAVID G., *An Approach to Measuring Results in Social Work* . . . . . JOHN E. WALSH 408
RANDOM DIGITS 410

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

The Editors welcome the submission of manuscripts for possible publication. They should be typewritten entirely double-spaced, including footnotes, and two copies should be sent to the Editor, W. Allen Wallis, 207 Haskell Hall, University of Chicago, Chicago 37. Books for review should be sent to the same address. Unsolicited book reviews are not accepted, but suggestions of titles for review are welcome.

EDITOR

W. Allen Wallis, *University of Chicago*
Assistant to the Editor: Margaret A. Labadie

ASSOCIATE EDITORS

Harold A. Freeman
Massachusetts Institute of Tech.
George M. Kuznets
University of California
Philip J. McCarthy
Cornell University
I. Richard Savage
National Bureau of Standards
C. Ashley Wright
Standard Oil Company (N.J.)

ADVISORY PANEL OF FORMER EDITORS

William G. Cochran (1945–50)
Johns Hopkins University
William F. Ogburn (1920–1925)
University of Chicago
Frank A. Ross (1926–34, 41–45)
Thetford, Vermont
Frederick F. Stephan (1935–40)
Princeton University

Errata: Readers and authors are urged to submit to the Editor notices of errors found in this or any previous volume. These will be published once a year, in the December issue.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 266 JUNE 1954 *Volume 49*

MEASUREMENT FOR ECONOMIC MODELS*

STANLEY LEBERGOTT
Bureau of the Budget

Three chief subjects are considered. (1) How accurate should data be for economic analysis via economic models, particularly input-output models? Preference is expressed for measurement of additional aspects of economic phenomena rather than broadside improvements in generally sound series. (2) How can models be developed to utilize imperfect data? Several suggestions are made: (a) examine the concepts and methods lying behind the basic data; (b) question models acutely sensitive to the inclusion of a single observation; (c) evaluate the economic meaning of models after empirical testing; (d) set up check models. (3) How can more adequate data be developed? A proposal is outlined for an integrated set of data on financial aspects of business, and for similar integration of data on employment and on consumer economic behavior.

MORE than half a century has passed since Marshall first sketched a system of equations which could be developed "until they embraced within themselves the whole of the demand side of the problem of distribution."[1] In that period the development of systems of equations, or models, has become a commonplace in economic study. Have our measurements kept pace with the demands which such systems place upon them? Three central questions can be set out for consideration:

1. How accurate should economic data be for model building?
2. How can models be developed to utilize imperfect data?
3. How can more adequate data be secured?

* A paper presented at the annual meeting of the Econometric Society, December, 1952. The opinions expressed are not necessarily those of the Bureau of the Budget.

[1] *Principles*, (London, Macmillan, 1916) Mathematical Appendix, Note xiv

I

How accurate should economic statistics be for model building? Clearly we do not wish them to be as accurate as possible: no one is prepared to devote the time and money required to achieve this end. We could, for example, immediately improve our estimates for gross national product, gross private domestic investment, new construction, personal consumption expenditures, incomes of unincorporated enterprise, personal savings, net realized farm income, index of prices paid by farmers, parity income index, and a host of other series if we had substantially better data on how much farmers spend for hammers and nails. We could reduce apparent inconsistencies between three sets of data on food consumption—Department of Agriculture estimates of food disappearance, the food component of the gross national product series, and the Census of Business data on food sales at retail—if we had a comprehensive study of food consumption outside private households. These projects, however, certainly lack the glamor and usefulness of many other projects; and it is most unlikely that funds could be secured to carry them out.

For general economic analysis, whose questions have been with us for a good many years, it is clearly difficult to demonstrate how accurate statistics should be. Nearly half a century ago, for example, Alexander Dana Noyes analyzed a period of recession, using statistics which reported a 50% decline in iron output, a 25% fall in textile output, a 55% rise in commercial failures and a 12% drop in railway receipts.[2] In 1886 Carroll Wright reviewed the course of the 1882–85 depression, noting the proportion of establishments idle and the number of business failures reported by R. G. Dun, commenting on steel and coal production series, on savings bank deposits and price trends, and concluding that the depression caused a "crippling of the consuming power of the people" but that "the volume of business transacted is not crippled comparatively to any such extent."[3]

Since these early days there have been obvious and welcome improvements in the accuracy of our data on industrial production and prices—to mention but a few. Yet if we review the reports in business magazines, the annual reviews of the Council of Economic Advisers and the Congressional Joint Committee on the Economic Report, can we state with assurance that such improvements have made any sig-

[2] A. D. Noyes, *Forty Years of American Finance* (New York, Putnam, 1909), p. 377–78.

[3] First Annual Report of the Commissioner of Labor, March 1886, *Industrial Depressions* (1886), p. 70.

nificant difference in the ability with which businessmen or government officials can evaluate economic trends?

Is the issue any clearer for model builders? Professor Morgenstern has pointed out as "obvious" that in large scale numerical operations such as those involved in inter-industry analysis "only very high quality data" should be used—given the extent of the numerical operations and their considerable cost.[4]

In actual practice how should such a prescription be applied? One of the dubious rows in the comprehensive and invaluable 190 by 190 inter-industry matrix which the Bureau of Labor Statistics recently prepared is that which distributed the output of "other repair services." The row for retail trade is only somewhat less dubious. Statisticians in the BLS and elsewhere have clearly recognized the inadequacies of these data and insisted on the desirability of improving them. Yet in the actual use of the model can we expect different policies to be pursued, or even suggested, if the aggregate demand for "other repair services" or for trade services differs by 1%—by 10%—by 20%? Probably not. On the other hand variations of 10% in the projections for copper demand might well be significant. Proceeding further, does this mean that all the coefficients indicating demands for copper by particular industries should be improved? Not necessarily. Of an estimated $1.3 billions in gross domestic output of copper in 1947, only some $200,000 was used by the industrial inorganic chemical industry. Clearly it is more important in terms of the final use of the matrix to improve not the coefficient for copper use by this industry but rather for use by the insulating wire industry, which consumed $345 millions—or even better, by the new construction industry which, on the crudest kind of data, was estimated to consume $162 millions.

Let us look at this from another viewpoint. As the size of the chart grows the size of the average coefficient diminishes—as does the likelihood that the error in estimating any given coefficient will significantly distort the estimates of output which are derived from models using the matrix. Hence the case for broadside improvements of data decreases in potency.[5]

It is particularly unfortunate that so many important studies of the

---

[4] Oskar Morgenstern, *On the Accuracy of Economic Observations* (Princeton, Princeton University Press, 1950), p. 39.

[5] Offsetting errors within a closed system will, of course, further minimise errors of estimate. However, there is no clear assurance that they will offset within the area of particular interest—e.g. demand for a particular industry's products—and *a fortiori* this argument does not apply in the usual open system.

stability of input-output coefficients, of the accuracy with which projections are made by models using input-output matrices, should have emphasized average errors or across-the-board accuracy.[6] Let us concede that it is useful to possess such overall measures. At the same time we must note that these measures are unrealistic—unrealistically harsh on input-output models as well as unrealistically easy on them. They are overly harsh because a complete improvement of all coefficients is certainly not to be achieved: it would cost as much as dredging a yacht basin or similar projects to which a distinguished senatorial economist has called attention.

Yet at the same time they are too generous insofar as they permit the implication that errors in estimating the total output of steel are no more forbidding than equal errors in estimating the total output, say, of trade. Clearly the difference is critical if such projections are not designed as mere mathematical exercises but intended for potential policy use.

Harold Barnett's valuable study makes it possible to empasize this distinction.[7] From his data we can contrast how the errors made in projecting the actual 1950 indexes of output for all industry groups compare with those for subtotals of separate industry groups. Let us set up two categories—one to include industries for which input-output projections would be of lesser interest since they would probably not be limiting elements in a full employment or mobilization situation—e.g. agriculture, food processing, etc.[8] The second group would include chemicals, metals transport, and the more limiting industries. In the consumption model the average errors in projecting 1950 indices by the 1939 input-output matrix were as follows:

All industries:	31.5	index points
"Less limiting" industries:	23	index points
"More limiting" industries:	36	index points

In the investment model the differences were even more marked; errors for the more critical industries were 41, or almost double the 21 for the less critical industries.

---

[6] Cf. *inter alia* Selma Arrow, *Comparisons of Input-Output and Alternative Projections, 1929–39* (The Rand Corporation, 1951).

W. W. Leontief, *The Structure of the American Economy, 1919–39* (New York, Oxford University Press 1951), pp. 216–18. These figures have been superseded by a comparison for 1950 outputs using the 1939 matrix in an unpublished memorandum by Professor Leontief and J. Fei. To some extent Waugh's analysis as applied by Christ is an exception. Cf. Part III of Carl F. Christ, "A Review of Input-Output Analysis," in *Conference on Business Cycles* (New York, National Bureau of Economic Research, 1951).

[7] Harold Barnett, *Specific Industry Output Projections* (The Rand Corporation, 1951), Table 4 gives deviation in index points from actual 1950 output levels.

[8] Agriculture, food processing, lumber, furniture, wood and paper, printing, textile, apparel and leather.

The inference I would draw from all this with respect to the improvement of data for input-output and similar models is that we attempt increasingly to separate the sheep from the goats: industries with reasonably well behaved coefficients where erroneous projections of their output would involve small losses should be estimated as accurately as possible once—and thenceforth given second place in our attention relative to a very limited number of industries (such as steel foundries, primary copper, etc.) and the vast area of guessing the exogenous variables. To some fair extent this is what has been done in empirical research on input-output coefficients. We clearly do not require the mammoth sample survey of all manufacturing industries which has been proposed in the past in order to produce coefficients for all industries of "a very high quality."

All of which still does not solve our initial question. We may agree that neither improving the accuracy of all economic data nor achieving maximum accuracy for most data is a practical aim. But how accurate should any such data be? The question, I think, can only be solved *in ambulando*. There is no easy way of calculating the probable loss which results from errors in estimating economic data of the kind used by most econometricians. This is true because policy decisions do not usually rest on a single datum, a single series nor a single model. And it is even truer because errors in the choice of a model can have a far more critical impact than errors in the data.[9] It is up to the data producers, model builders, and data users generally to express their judgments on what levels of error are so intolerable that they believe more money should be spent for improving the data.

I should like to express a general preference for additional measurement series rather than improvements in existing series or data. Let us grant the obvious exceptions and assume, moreover, that the data are collected by persons with at least a minimum budget and minimum statistical competence. Then I think it can safely be said that knowledge about additional aspects of our economy will generally make a greater contribution to sound theory and sound policy than will improvements in the accuracy of existing series.

As one instance let me cite the patient and ingenious labors which produced Fabricant's index of manufacturing output.[10] The increases in accurate detail have been invaluable. Yet so far as concerns the totals for industrial production—and it is these which are used so

[9] The post-war forecasts erred because of the choice of model, the choice of consumption functions, because of ignorance about the way the labor market functions, business plans, business reaction to tax schedules and so on. Errors in the data were of small moment.

[10] Solomon Fabricant, *The Output of Manufacturing Industries, 1899–1937*. (New York, National Bureau of Economic Research (1940).

widely in economic models and policy decisions—the resultant index differs little from the one which Day and Thomas developed from far worse data with far less work.[11] Furthermore, neither index for total manufacturing would differ greatly in movement over the years from a very crude combination of output indices for textiles and iron and steel.

A similar consideration is apparent when one compares the recently revised and unrevised FRB production indexes, wholesale price indexes, etc.[12]

The substantial differences in series occur when a responsible investigator with a reasonable minimum of resources, cooperation and time first develops a series on a sound conceptual basis. Further improvements may be real; they certainly are hard won and usually expensive. They usually profit the model builder very little.[13]

In pressing for improved measurement in economic statistics the model builder would be well advised to give increased attention to the basic opportunity cost question. That question is this: do we improve our understanding of a particular economic phenomenon more by improving still further a particular series (or datum) or by adding new series on yet uncharted aspects of the phenomenon?

II

Given the present existence, and probable persistence, of errors in economic observation what is the model builder to do? How is he to build models for utilizing imperfect data? Models allowing for a random run of disturbances have been developed—well developed—in recent years and there is hardly need to recommend their use to econometricians. But if we are concerned with biased data, where acute errors are possible—and those not randomly disturbed—some alternative action is required.

A. The first requisite for the model builder is to examine the concepts and broad methods of estimation used in deriving the data which he incorporates into his model. I realize that this may be considered an undue burden; developing a sound model is a more than sufficient labor. Yet the alternative is even more unsatisfactory.

---

[11] E. E. Day and W. Thomas, *The Growth of Manufactures, 1899 to 1923* (Washington, Government Printing Office (1928), pp. 34, 94. With 1939 = 100, the 1899 estimates are 28 for Fabricant and (32) for Day-Thomas. Succeeding estimates are 34(39), 43(51), 51(55), 61(69), 54(54), 77(85), 82(88), and 87(88).

[12] The examples may be new, but the observation is an old one. Cf. Wesley Mitchell, *History of Prices During the War* (Washington, Government Printing Office, 1919), p. 28, where he indicates the similarity between the BLS index, based on 350 commodities, and the War Industries Board index based on 1,474 commodities.

[13] None of this, of course, bears on the question of what topics happen to evoke the abilities of research workers: if first class statisticians are willing to toil for years in the statistical salt mines, theirs is—and should be—the choice of mines.

Let us consider, for example, the ill charted and forbidding area of consumption projections. Klein is outspoken in asserting that "to many of us engaged in econometric work it became obvious in the second half of 1947 that the most serious deficiencies in the existing models lay in the consumption equation and in the group of relations serving to determine absolute prices."[14] Despite the wealth of data available on the subject, the relations between consumption and income continue to baffle projectors. On the one hand the budget surveys uniformly report that the proportion of income saved is greater at upper income levels than at lower ones. On the other hand Kuznets' data on capital formation seem to indicate that the aggregate consumption function has not changed perceptibly in the past sixty years when incomes have risen so.

In an acute discussion of the factors affecting consumption Professor Fellner has emphasized that

> The most obvious characteristic of the historical consumption function, as calculated from Professor Kuznet's estimates, is that it does not show a tendency to flatten out. It tends to linearity regardless of the varying population growth of the subsequent historical periods.[15]

A recent path breaking study on consumer behavior similarly notes that "the Kuznets' data do not show any trend in the savings ratio," suggesting that even if their level were incorrect they "allow us to make a judgment about the movement of the savings ratio."[16]

Reference to the basic sources indicates that to a very great extent this constancy exists because it was estimated that way. (And, incidentally, it was estimated that way because a sensible model indicated that to be the best method of estimate.)

More specifically: the consumption function relates consumption to net product. The latter consists of two segments: consumption and investment. Obviously consumption correlates perfectly with itself. But the investment segment also will reveal a high correlation with consumption because of the method of estimate. Let us consider in turn each of the components of the investment total.

a) *Inventory change.* Changes in manufacturing, trade and "all other" inventories—together accounting for more than half the total inventory change in most years—are estimated by constant ratios to output.[17]

---

[14] *Conference on Business Cycles* (New York, National Bureau of Economic Research 1951), p. 117.

[15] William J. Fellner, *Monetary Policies and Full Employment* (Berkeley, University of California Press, 1947) p. 56. Cf. also Paul Samuelson, in Seymour Harris ed. *Postwar Economic Problems* (New York, McGraw Hill, 1943), p. 33.

[16] James Duesenberry, *Income, Saving and the Theory of Consumer Behavior* (Cambridge, Harvard University Press, 1949), p. 56.

[17] Simon Kuznets, *National Product Since 1869* (New York, National Bureau of Economic Research, 1946), pp. 109, 110.

Now since a heavy proportion of total output is output of consumer goods, this method of estimate means that these inventories will necessarily fluctuate with consumer expenditures. Changes in inventories of livestock represent a further link, since livestock slaughter (a component of consumer expenditures) was estimated, in the original Department of Agriculture source, by applying ratios to inventories.[18]

b) *Construction.* The output of construction materials was used (with a weight of one, plus semi-durables with a weight of two) to interpolate estimates of consumer durable output between census dates.[19] This is a further link between the consumer and producer segment estimates.

c) *Producer durables.* The basic trends for producer, as for consumer durables, derive from the 1869, 1879, and other Census date totals estimated in William Shaw's comprehensive and fundamental study.[20] For most durable categories Shaw extrapolated 1914 and 1909 totals to earlier years by using a constant ratio to split Census group totals between consumer and producer durables.[21] The procedure assumes, for example, a constant ratio of consumer to producer goods in the output of furniture, sewing machines, foundry and machine shop products, etc. Thus 67.7% of all sewing machines were assumed to be household machines in every Census year from 1869 through 1909, with the balance classed as business use.[22] Other constant ratios (under 100%) were used for furniture, heating and cooking apparatus, appliances, cutlery, etc. For an important class of items Shaw made the reasonable assumption that none were bought by producers, all by consumers, at each of the Census dates. These include "family and pleasure" carriages and wagons, pianos, organs, rugs, glassware, books, etc.[23]

So far as concerns the level of commodity flow results, or even many broad economic trends these decisions raise no question: a responsible estimator cannot arbitrarily vary ratios without some data to which to tie. But if we are concerned with something as precise as the constancy of the historical consumption function, we should decide how critical was the decision to assume that the rising number of apparel firms did not affect the proportion of all sewing machines bought by business, or the assumption that the growing number of restaurants and hotels did not change the proportion of all dishes and cutlery they bought,

---

[18] F. Strauss and L. Bean, *Gross Farm Income and Indices of Farm Production and Prices in the United States, 1869–1937*, (Washington, Government Printing Office, 1940), p. 106.

[19] *Ibid.*, p. 95.

[20] William H. Shaw, *Value of Commodity Output Since 1869* (New York, National Bureau of Economic Research, 1947)

[21] *Ibid.*, Table III.

[22] William N. Shaw, *op. cit.*, p. 161. For 1879–1909, 3.1% of the reported total is excluded as unfinished.

[23] *Ibid.*, Table III and Note B to Table III.

etc. When we come to use the data in economic models, we must not try to explain any constancy of savings ratios in final series if it rests ultimately on the constancy of arithmetic ratios used in the estimating process.

B. In addition to a careful examination of the methods of estimate a second step is to question models which are acutely sensitive to the addition or subtraction of a single observation—such as that for a single year or a single industry.[24]

A recent model which appears to be in this category is Modigliani's savings model.[25] Modigliani defines consumer saving as a function of the current year's income and highest previous year's income as measured by a cyclical income index. The model provides a beautiful fit for the 1921–40 period. If, however, we simply add the values for 1941, as provided in the appendix to Modigliani's study, we derive an estimated 1941 value so far in error that it cannot even be plotted on the chart he shows.[26]

His particular equation for allowing for the influence of previous years' incomes essentially states this—that for the entire decade of the 1930's the consumers continued to hark back to the halcyon 1929 level of incomes in deciding how much money to spend. However, the same model indicates that the consumer's frame of reference—say for deciding on his 1937 saving—would swiftly move forward six years from 1929 to 1936 if the 1936 total were a mere 4% greater.[27] (It would incidentally, take a courageous man to assert that Commerce 1936 income estimate is accurate to within 4%).

A fairly simple method to keep from implying that such drastic changes in consumer behavior occur when a single observation is changed slightly or added to previous ones would be to drop what is essentially a function based on ranking in favor of one based on a direct measure of magnitude. The latter would be less at the mercy of minor variations in source reporting or estimating techniques.[28]

Apparently, the basic equation for the demand for labor in Christ's

---

[24] Carl Christ wisely points out that if there were really a sharp change in important economic relationships then a model might properly be acutely sensitive to the addition of post-change data. One might add that economic relationships generally change so slowly that we should question both the data and the model, before we conclude that a sharp change in relationship has, in fact, occurred.

[25] Franco Modigliani, "Fluctuations in the Saving-Income Ratio: A Problem in Economic Forecasting" in *Studies in Income and Wealth*, (New York, National Bureau of Economic Research, 1949), Vol. XI.

[26] Franco Modigliani, *op. cit.* p. 380.

[27] The only consolation indicated is that since 1941 income levels were above any prior year 1929–40 the gain is defined as being a secular, and not a cyclical one. Consider the economic meaning of this. If per capita income rises $100, this is $100 of cyclic increase; if it rises $101, none of it is cyclical.

[28] A distributed lag approach was wisely suggested by Leontief; Modigliani noted that he had tried this and secured worse results. But less precise correlations based on due recognition of data shortcomings are to be preferred to the converse.

limited information version of the important Klein model III is almost equally delicate. By adding two observations to the 21 year run from 1921 through 1941—which had included short recessions, deep depression, and continued prosperity—Christ obtains a projected 1948 demand for labor which Klein reservedly calls "fantastic"—a comment echoed by Christ.[29] This, moreover, is no trivial or paltry incidental equation: it is fundamental to the model. It seems a striking fact that the disturbances which were trivial and random in Klein's limited information estimates become substantial and nonrandom in Christ's versions: they are negative for 1924, 1927, 1930, and other recession years, but positive (with one exception) for all other years.[30]

Christ assumes that the problem lies in the limited information method at least as much as with the equation. There is merit in this argument—but not enough. While the limited information solution actually blows up the standard error of disturbances even for the least squares solution is sufficiently disturbing—rising from .96 to 1.47 when Christ adds these few years to the original run. Klein favors errors in the basic data—particularly pointing to inadequacies in the BLS consumers price index. While the price data are dubious this is only part of the answer.[31] It would appear that the chief difficulty lies in the fact that Christ's calculations imply only a $9 billion rise in consumer expenditures from 1946 to 1947, whereas the Commerce data (published after his original work) report on $17 billion rise—reason enough to show the inconsistency between output and employment data which he and Klein discuss.[32] What is more important than the particular reason is the fact that any system which is so sensitive to the addition of several observations would seem to fall into the "handle with care" class.

C. There is a third consideration for model builders to bear in mind when using data—particularly imperfect data. And that is the need to evaluate the economic meaning of econometric models after empirical testing. Just as we will accept a sampling procedure which gives biased results provided the total errors are considered satisfactory so

---

[29] Carl Christ, "A Test of an Econometric Model for the United States, 1921–1947" in *Conference on Business Cycles* (New York, National Bureau of Economic Research, 1951), p. 124.

[30] *Ibid.*, p. 105.

[31] Using the BLS and BAE sources which Christ notes (*ibid.*, p. 91) one can estimate his price rise from 1946 to 1947 at 16%. This may be compared with the 10% gain shown by the deflator given the 1951 *National Income Supplement*, p. 146.

[32] Taking Christ's consumption data from his study (*ibid.*, p. 90) and applying the 16% rise computed above indicates a rise in consumption in current prices of $9 billions—as compared with a rise of $17 billions in the comparable Commerce consumption expenditure totals, (exclusive of imputed rent.) For the discussion by Klein and Christ, see *ibid.*, pp 98–101 and 114–15 especially. The error apparently arose in the use of regression procedures to extrapolate the Commerce figures.

we should prefer a model with greater resemblance to the economic world instead of one that has somewhat less error and distinctly less meaning.

It is instructive on this point to consider the recent study in which Tobin sets up a sensible model of the food market, combining budget data and time series data with equal parts of care, agility and aplomb.[33]

Tobin posits the following model of the retail food market:[34]

$$S_t = KY_t'^{\alpha_1} Y_{t-1}'^{\alpha_2} P_t^{\beta} Q_t^{\gamma}$$

where S^t is per capita food supply for domestic consumption, Y^1 is disposable income, P represents food prices and Q, all other prices.

Converting to a reduced form and rewriting, he developed his system as follows:

$$\log P_t = b_0 + b_1(\log S_t - \alpha \log Y_t') + b_2(\log Y_t' - \log Y_{t-1}') + b_3 \log Q_t.$$

On a subsequent page, Tobin computes values for the parameters using various Agriculture and Commerce Department series for 1913–41. The values for each "b" are negative. The economic meaning of Tobin's model then seems to be this: When disposable income rises food prices tend to fall; and when food consumption rises more than past income-elasticity relationships would have predicted there will likewise be a tendency for food prices to decline.

This fairly odd pair of inferences suggests, I think, the desirability of evaluating the economic meaning of the model and its empirical parameters in some detail. The fact that the food consumption series is far less sensitive than the series for income or prices may offer an explanation. Its range is a mere 10% over the period Tobin uses it—whereas food prices and disposable income each had a 100% range.[35] This lack of sensitivity of the food consumption index suggests that some alternative measure of food consumption might properly be used —e.g. the deflated food expenditures figures of the Department of Commerce. A detailed comparison between the BAE and the Commerce indices (even after adjusting for such conceptual differences at the Commerce inclusion of services) strongly suggests that the two series do not move in consistent fashion.[36]

---

[33] James Tobin, "A Statistical Demand Function for Food in the U.S.A.," *Journal of the Royal Statistical Society*, Vol. CXIII, Part II, 1950.

[34] *Ibid.*, p. 130.

[35] This may well raise the question whether an identification problem is involved. Tobin carefully examines this question (*Ibid.*, pp. 135–6) and contends that it is not.

[36] Tobin infers that the failure to include in the series services in food distribution, would understate changes "in the supply of 'finished' foodstuffs." But nevertheless uses it—presumably as being reasonably sound. *Ibid.*, p. 131.

These are essentially negative cautions—though evaluating the methods by which his data were originally secured seems an elementary caution for any model builder to take, however pedestrian and however unrewarding such an inquiry proves to be. And analysis of the economic meaning of the resultant model—after the coefficients have been determined—seems equally urgent.

d) Positively, however, there is something which econometricians can do until that happy day when all economic data come from a single well designed survey producing data of only the highest reliability What can be done is to set up a check model—or models—to parallel the basic model. Such a model would incorporate related series or measurements in place of those used in their basic model. Clearly one would hardly seek to redo the 650 equations of the BLS 1947 matrix. Nor is this necessarily required. What we do need is a truncated model to give us some feeling for what changed inferences from the model would be produced by given errors in the original data. Let me be specific.

In recent models some fairly complex allowances have been made for changes in the distribution of income. To demonstrate that any of these allowances are real advances over, say, the simple procedure used by Tinbergen and Kalecki of distinguishing wage from non-wage income, we should have some measure of the adequacy of reported income distributions. One method would be to compare approximately similar distributions. Table I gives two distributions—one for 1934–35 and one for 1935–36.[37] The apparent differences between the 1934–35 and 1935–36 distributions are great—far greater, in fact, than the differences between surveys of slightly different populations separated by four years (1935–36 to 1939) of rising income.[38] A similar comparison for 1949, say, can be made between the Census and FRB surveys, and the differences are much smaller. In both instances the model builder might usefully use first one, then the other distribution to get an indication (*not* a precise measure) of the impact of survey variations on his conclusions.

A second method would be to make arbitrary adjustments in the

---

[37] Faith Williams and Alice Hanson, *Money Disbursements of Wage Earners and Clerical Workers in Five Cities in the West North Central-Mountain Region, 1934–36*, BLS Bulletin No. 641 (1939), p. 114. The concentration of the schedules in 1934–35 is noted on p. 341; A. D. H. Kaplan, *et al.*, *Family Income in Seven Urban Communities of the West-Central-Rocky Mountain Region*, 1935–36, BLS Bulletin 646, No. 1, p. 125.

[38] 1935–36 data from BLS Bulletin 646, No. 1, p. 110. 1939 data from 1940 Census, *Families, General Characteristics*, Table 59. These data relate to all families receiving wage or salary income. A distribution for all families without income other than wages & salaries would be virtually identical—for the $500 and over group discussed here.

data on the assumption of specified margins of error. For example, the income data reported in the BLS survey of 1941 aggregated about 10% less than control totals indicate was received.[39] More recent surveys have run from 5 to 10% low. What happens to our models if we allow for such understatements at each date, or if we presume the possibility that the underreporting of income decreased from the 1901 to the 1949 expenditure surveys?

TABLE I

DENVER WAGE EARNER AND CLERICAL WORKER FAMILIES WITH INCOMES OF $500 AND OVER PER CENT DISTRIBUTION BY INCOME LEVEL

	White nonrelief families*		All Families†	
	1934–35	1935–36	1935–36	1939
$ 500– 999‡	14	18	30	24
1,000–1,499	35	29	28	23
1,500–1,999	36	24	19	19
2,000–2,499	13	15	11	13
2,500–2,999	2	7	6	8
3,000 and over	0§	7	6	13
$ 500 and over	100	100	100	100

* Complete families, husband and wife native born.

† Families receiving $500 or more in wage or salary income.

‡ Families with incomes under $500 excluded from 1934 survey; also excluded from 1935–36 and 1939 data here for comparability.

§ Less than 1%

The measures need not be as close conceptually as this and cannot be in most instances. However, trends in the number of nonfarm employees (as measured by the Census Bureau's Current Population Survey) and in production (as measured by the FRB industrial production index) should be definably similar. Trends in manufacturers sales as reported by the Commerce sales series, and the FRB production index times the BLS price indexes should also bear a close relationship to one another and so on. The analyst can assess the result of using variant measures, deciding how much appears to be accounted for by conceptual differences, how much results from the fact that he chose one series rather than another.

[39] Selma F. Goldsmith, "Appraisal of Basic Data Available for Constructing Income Size Distributions," *Studies in Income and Wealth*, (New York, National Bureau of Economic Research, 1951), Vol. XIII, p. 285.

An additional reason for the use of check models is that we inevitably concentrate on one slice of reality, one aspect of the phenomenon being examined when we select any given series as an explanatory variable. It is frequently not clear whether that particular aspect is the most relevant, real or useful one to consider. Designating alternative measures helps triangulate the area of concern. As E. A. Goldenweiser remarked some years ago:

> no set of statistical series, to say nothing of any single series, is a sufficient basis for determining casual relationships on which economic policy can be predicted with safety. They are only indications of where one should look for the causes and interrelationships that determine economic events.[40]

The moral applies as well to the use of models for analysis.

III

How can we secure more adequate statistics both for policy determination and for economic analysis by economic models?[41] The simple answer would be to pass a law—not on statistics, of course, but on matters requiring statistics for their determination. Certainly, for example, the operations of the Securities and Exchange Act have developed better data on financial flows than even the combined talents of Crum, Epstein and others had been able to produce by the process of making bricks from straw. The passage of the Agricultural Adjustment Act brought support for agricultural income statistics of a kind that was rare even in the Department of Agriculture.

A more realistic answer to the question, however, must start from an answer to the question: what kind of data are needed for economic models? The answer, I believe, is essentially the same type of data as is required for general economic analysis except that it is much more urgent that the data be consistent. It is more of a problem for the model builder than for the general economist that the official series on plant and equipment expenditures is not consistent with that for sales in machinery and related industries; that one government series can report a 7% fall in auto inventories while another reports a 0.2% rise for the same period; that one series reports a drop in profits from

[40] "The Economist and the State," *American Economic Review*, March 1947, P. 5, Cf. also Charles D. Stewart and Loring Wood, "Employment Statistics in the Planning of a Full Employment Program," *Journal of the American Statistical Association* (September 1946).

[41] These purposes are not exclusive. In the preparation of the Federal Budget a set of economic projections are made for national income, prices, employment and other materials organized as a model. Similar materials are prepared by the staff of the Joint Committee on the Economic Report, Cf. Samuel M. Cohn, "Managing the Expenditure Side of the Federal Budget," a paper presented before the American Society for Public Administration, Washington Chapter, November 7, 1952. Joint Committee on the Economic Report, 82nd Congress, 1st Session, (1951) *The Economic and Political Hazards of an Inflationary Defense Economy*, Cf., especially Appendix C.

food manufacturing of 12% while another equally reputable series reports a rise of 20% that comprehensive reports on manufacturing employment for the same year may differ by several million depending on which series is selected; that employment data will be plant reports, production data may be company reports and profits data from consolidated company reports—each of these differences limiting our use of all three variables in model building, and so on through a catalog that is long, but not quite everchanging. To achieve such consistency, plus an improvement in accuracy, the following program areas deserve consideration.[42]

1. *Financial aspects of business.* Basic benchmark data can be provided annually for all of business, corporate and unincorporated from tabulations of the corporate, partnership and individual proprietorship income tax returns to the Bureau of Internal Revenue. Basic current data, used to extrapolate the benchmark data from quarter to quarter, can be secured from an expansion of the Financial Reporting Program now conducted jointly by the Federal Trade Commission and the Securities and Exchange Commission. That program, now restricted to manufacturing corporations, should be expanded to other industries, to unincorporated business. Now securing data on profit and loss, on balance sheet items, it could reasonably secure related data on business investment orders and plans for future sales and investment.[43]

The stakes for analysis here are first, accuracy; second, consistency; and third, a considerable increase in the number of observations. Instead of being restricted to one annual value for a year, only imperfectly reflecting a change in economic direction, the analyst could look forward to from 4 to 12 observations a year for a linked set of variables. The result should bring a much better knowledge of economic variation and relationship.

2. *Employment.* In the models developed by Hagen, the NPA, Fortune and other postwar projectors, in Klein's model III, and in the recent emergency model sponsored by the Air Force, production data, establishment reported employment data (from BLS), and household

[42] Richard Stone has argued persuasively for a single sample survey for business, one for consumers and one for government. Cf. his *The Role of Measurement in Economics* (Cambridge, University Press, 1951), pp. 57 ff.; R. Stone, J. E. G. Utting, and J. Durbin, "The Use of Sampling Methods in National Income Statistics and Social Accounting," in *Revue de L'Institut International de Statistique*, Vol. 18, No. 1–2 (1950); p. 31; and corresponding comments in reports of the UN Sub-Commission on Statistical Sampling. The consolidated surveys proposed here attempt to achieve the same goal of consistent data but at the same time make use of the mass of reliable data made available regularly in the U. S. from administrative reports.

[43] Administrative relationships are not particularly relevant here but it might be noted that in practice such a program would make use of such existing collection mechanisms as the Census Bureau's Current Business Reporting sample, etc.

reported employment data (to the Census Bureau) are all utilized. With labor a limiting resource in some projections, and unemployment an ominous resultant in others, the consistency of these data is particularly important.

Benchmark data for the basic BLS series come from the State Unemployment Insurance and Old Age and Survivors Insurance Systems. The reports to OASI, however, are simply tax reports to the Internal Revenue Service. It is worth considering whether the system of reports to the IRS might be so integrated that business would report annual payrolls and employment on their income tax returns.[44]

If this were done we could then look to the monthly Bureau of Labor Statistics reports, like the FTC-SEC reports, as a means of keeping up to date on changes—with the additional consideration that the BLS data might be sought as in the past to apply to individual establishments rather than (as in BIR) to entire firms.

Such an integration of date on employment sales etc. would still exclude the Census Bureau's Current Population Survey. One approach would be to treat the list of establishments represented by workers who fall into the Current Population Survey as the sample for use by the Bureau of Labor Statistics.[45] By regular sample checks against establishment records of reports by individuals as to their employment status we have a method of reducing sampling differences between the BLS and Census surveys and removing characteristic errors in the responses secured by the population survey.[46]

3. *Consumers financial activity.* Moving from the sphere of business activity to that of consumers we are confronted with similar considerations. The analyst at present is likely to secure consumption functions from the well established BLS surveys. These functions may or may not be consistent with asset-savings relationships derived from the FRB Surveys of Consumer Finances, employment-income relationships estimated from the annual Census surveys or income elasticities for the consumption of food items as derived from data secured in

[44] BIR income tax returns do not now report employment, and permit the distribution of pay rolls under at least two deduction items.

[45] This proposal is discussed at greater length in the writer's "Labor Force Statistics: The Task Ahead," a paper presented at the 1950 annual meeting of the American Statistical Association.

[46] As has been emphasized by specialists in this field, errors in reports by the establishments would also have to be minimized—though this is presumably part of the regular program of the BLS. Administrative procedures would preclude sample changes more frequent than every year or so. However, if the survey is viewed as basically a means of interpolating between benchmark employment totals such an interval creates no real problem. The rate at which the sample should be changed must be judged in relation to the present rate of attrition and nonresponse in employment reporting by establishments. Accounting for the varying probabilities with which establishments would fall into the sample—i.e. what is the residence distribution of their employees; what is the correlation between employment characteristics of persons in the same family—is one of important technical issues which would have to be solved before such a proposition could be developed further.

Bureau of Human Nutrition and Home Economics and private surveys made for the Department of Agriculture. The survey data in turn are not fully consistent with the aggregates on income, saving and consumption that have been derived from business records and embodied in the national income accounts.

A program in this field to be of value for economic models must provide consistent data, minimizing the number of measurements coming from completely different surveys. To achieve this purpose use can be made of the Census Bureau's Current Population Survey for securing an integrated body of data on incomes, expenditures, asset holdings, family employment status and demographic characteristics, as well as price and income expectations.[47] The gradual development of such a program would remove the considerable inconsistencies which now develop when the model builder incorporates measurements for many of these factors from the variety of surveys which now provide them.

A program in this field must provide reliable data, and if possible provide measurements which reflect variations in consumer behavior as variations in economic activity occur. In practice this would mean securing monthly reports from consumers on their economic activity in the previous month or week.[48] Such reports would facilitate accurate reporting by consumers, since the memory feats required would be minimized. No less important is the likelihood that their frequency would give us an improved knowledge of consumer activity. For example, interviews for the *1948 Survey of Consumer Finances* were conducted during January and February of 1948. Some 50 percent of the respondent interviewed in January expected a price rise. The grain market broke in February, and only 15 percent of those interviewed after the break expected prices to rise.[49] For the model builder the increasing number of observations at different levels of the nation's

---

[47] This proposal was discussed at greater length in a paper prepared for the 1949 annual meeting of the American Statistical Association, "The Validity of Interviews: Consumer Expenditure Surveys." To minimize respondent burden and nonresponse a system of replicated designs would be necessary: broad information would be taken from the main sample and subsidiary details from related samples.

Small scale studies by the Bureau of Human Nutrition and Home Economics suggest that some of the advantages of replicated samples may be delusive but that the procedure is practicable. Cf. Barbara B. Reagan and Evelyn Grossman, *Rural Levels of Living in Lee and Jones Counties, Mississippi, 1945, and a comparison of two methods of data collection* (October 1951) USDA Agriculture Information Bulletin 41, esp. Part 2.

[48] More than one period may be used, since reporting for long periods may provide reliable data for some classes of items (e.g., cars purchased), whereas briefer periods may be necessary for other items (cigarettes, milk, etc.).

[49] Federal Reserve Board of Governors, *1948 Survey of Consumer Finances*, Part I, Table 7. The relative change for farm operators and other groups as a whole was almost identical. Cf. James C. Davies, "Some Relations Between Events and Attitudes" *The American Political Science Review* (September 1952) p. 780

economic activity should increase the number of variables which he can include in his model, or test with any conclusiveness.[50]

Children's stories used to end with the magic words, "and they lived happily ever after." The prospect for econometricians, unhappily, is quite the reverse. Data will continue to be inadequate. What concepts lie behind the statistical measurements and what errors are hidden within them will continue to be issues for exploration. Econometricians will increasingly have to delay their more fascinating analytic work in order to ponder on the data and results, making certain that their findings do not merely quote errors, or assumptions inherent in the original measurements. They will in all probability find it essential to confirm each model with check models, making use of whatever scraps of evaluative information are offered them by data producers. In all this, however, there is at least one measure of consolation—the increasing frequency with which data producers feel obliged to check their data against control figures, to conduct post enumeration surveys, and to rely on scientific sampling so that at least measures of sampling errors will be available for data users. For both the producers of data and those who use them in economic models have a similar goal—to improve our understanding of economic change.

[50] It is, of course, possible to multiply observations even more cheaply, by a process of interpolation. This procedure produces useful data for some purposes. However, it does not really add to the number of independent observations of change through time. For a contrary view, in practice, Cf. Colin Clark's use of Barger's interpolated quarterly data in his "A System of Equations Explaining the United States Trade Cycle 1921 to 1941," *Econometrica* (April 1949).

TECHNICAL ASPECTS OF TRANSPORTATION FLOW DATA

R. Tynes Smith, III
Bureau of Transport Economics and Statistics
*Interstate Commerce Commission**

This paper discusses the technical aspects of two problems relating to the 1 per cent sample of rail carload waybills currently being secured by the Interstate Commerce Commission. One concerns the selection of the sample itself while the other involves the use of the sample information for estimating a desired transportation statistic.

It is proposed to show that the present waybill sample is a simple yet powerful tool for transportation analysis. After a short historical review of the background of the problem in general, the sample selection procedure is described. This yields adequate but biased results so an additional adjustment is made which eliminates this bias and gives an efficient and representative sample. After this description an example of the straight forward manner in which the waybill data may be used to get answers to complex transportation problems is given by describing the technique used to develop a series of rate indexes. Finally several methods by which the standard deviations of these estimated indexes were actually determined will be outlined. But before proceeding a brief description of a waybill may be helpful for those readers who are not transportation experts.

A bill of lading is first prepared by the shipper which tells the carrier, among other things, the kind and amount of commodity to be shipped, and the points of origin and destination. The carrier prepares a waybill from this document which accompanies the shipment and which also provides the basis for assessing the transportation charges. After the shipment is delivered, the waybill is audited by the terminating carrier to correct the charges and is then filed. This audited document, therefore, provides a record of the actual transportation services performed and the charges made for that service. A copy is illustrated on page 238.

The use of waybill data in the analysis of transportation problems is not new. As a matter of fact, results of waybill studies have been used in proceedings before the Commission for at least fifty years. The first nationwide study was made in 1932, and the present waybill sample currently being secured by the Commission represents the first attempt to get continuing information in this manner. There has been a gradual evolution both in the type of data collected and in the technique of collection during this period.

* This paper has not been considered by the Commission.

The first individual studies were made to develop specific information and often covered only the movement of single commodities or only the traffic between particular points. There was no sampling problem because all of the bills for all of the traffic involved in a specific period of time were selected. The data were usually presented simply as observed facts for that period.

The first nationwide study was made by the Federal Coordinator of Transportation based on waybills covering one day's terminations in the year 1932. However, it was clearly recognized at that time that this sample could not be considered as a representative one so only limited generalizations were made from the relationships developed. The material was tabulated and no attempt was made to expand it as an estimate of the total traffic for a longer period than the actual day covered.

The first major attempt to obtain a countrywide representative sample for a whole year was made by the Board of Investigation and Research. Each Class I railroad was requested by that Board to supply it with certain information for all carload traffic terminated on its line on one designated day in each month of the year 1939. Different days in the month were assigned to the railroads in any one region so as to allow a better sampling of brief seasonal movements and the dates for parallel roads were staggered. Staggering of dates was arranged so far as possible to catch every week day during the week twice during year. This design was a major advance because the resulting sample could be considered as representative of the total traffic for the whole year.

Several waybill studies were conducted by defense agencies during the war. The War Department started a continuous sample of the bills of lading issued for its commercial traffic within the United States. This was a systematic sample which at the beginning was secured by simply counting the documents and making a selection in the proportion desired. Refinements were made in this procedure from time to time and a more efficient method was finally developed making use of the terminating digits of the bill of lading number.

It was found that a number of important characteristics such as type of commodity were related to the originating point and so stratification by origin prior to sampling greatly increased the efficiency of the sample. The bills of lading were prenumbered serially in large blocks and furnished in groups to the issuing office so that the desired stratification could be accomplished by arranging the bills according to their numbers. Analysis of the systematic selection of these bills led to the choice of a combination of the terminating digits as a selection device

which made the initial arrangement of the bills by number unnecessary but which preserved the advantages of origin stratification.

In 1946 the Commission started work on its waybill sample design and, of course, gave consideration to past experience in this field. Two practical methods of selection had been developed, the first based on all traffic terminating on certain days, and the second based upon terminating digits in the waybill number. Each of these plans offered peculiar advantages and disadvantages.

The selection of all traffic terminating on specified days was favored by many of the carriers because of its apparent simplicity. They also felt that there was some advantage in doing the necessary work involved in a short period of time in order to get it over with. However, there were some serious objections to this type of sampling, especially in connection with the ability to secure adequate representation of short seasonal movements and with the difficulty of determining the possible sampling error involved in estimates prepared from the sample.

The difficulty in determining standard errors of estimates based on solid days studies lies not in the fact that these are cluster samples, but that the clusters (days) are usually chosen on a judgment rather than a probability basis. Even if a suitable method of probability selection were applied the efficiency would be low because the seasonal factor of much of the traffic makes the between days variance of many important characteristics quite high. Since practical sample sizes lie between four and twelve days for a year, it is evident that estimates would necessarily be subject to relatively large standard errors. The author was recently told of a case where a sample consisting of four days, one in each quarter, was found to contain no observations of an important commodity known to be moving in large volume. A check revealed that there actually had been no movement on the selected days due to chance circumstances such as strikes and floods.

An alternative method of sample selection is based upon the terminating digits of the waybill number. This was a new idea to most of the carriers and many of them felt that it would present a more difficult problem than would the solid day's sample. It was also determined that there were a number of railroads using monthly numbering systems which would produce a bias in this type of selection if a standard combination of terminating digits were used for selection. Balancing these objections were the very important considerations of greatly increased efficiency in the sample, including the ability to reflect very small seasonal changes, and the possibility of making reasonably accurate estimates of the sampling errors of sample results. After careful con-

sideration of these and other factors it was decided to base the sample selection on the terminating digits of the waybill number.

A waybill sample could be selected either at the point of origin or at destination but the terminated bill is preferred because it contains complete information regarding the actual movement of the shipment and audited results of the charges assessed. However, this choice presents a problem when the selection is based on waybill numbers because these numbers are prepared at the origin points. There are several common methods of numbering waybills and the possible effect on the sample of these methods must be considered in the sample design.

All waybill series are numbered progressively to some point and then start over again at the initial number. The end of the series may be determined by some period of time, such as the end of a month, or by some number in the series itself such as the waybill number 10,000, 100,000, and so on. This latter type of series may be called a block system and it can easily be seen that the selection of all waybills with numbers ending in any given pair of digits from such a series will yield a systematic and unbiased one per cent sample of the waybills if the length of the block is some multiple of 100. Also, since it appears that the traffic characteristics of a shipment are independent of the terminating digits of the waybill number on which it moves, an unbiased sample of waybills based on these digits will also be an unbiased and representative sample of the shipments. It is immaterial in this case whether the selection is made at origin or destination.

The practice of starting waybill numbering series over again at the end of a period of time, such as each month or year, presents a different problem. It does then make a substantial difference whether the sample selection is made at origin or destination. An example of the effect at several small stations numbering their bills on a monthly basis will illustrate.

Suppose a sample of one out a hundred waybills is desired from a group of stations each issuing less than a hundred waybills per month. If the sample is to be selected at the origin station, there are a number of acceptable devices which could be used. A simple and direct approach would be to establish a separate register for checking off the bills as they were issued which would indicate each hundredth one for selection. This method, of course, is independent of the type of numbering system or of the volume of bills issued. Other methods can be devised using the waybill number but they immediately run into the difficulty that this number alone is not sufficient to make a one per cent selection. There will only be as many different numbers as there are

bills issued each month so that in the case of small stations some method of compounding must be used if a one per cent probability of selection is required. It is evident that this compounding device must be based on the volume of traffic at the issuing station and the waybill number initially used cannot be greater than the total issued in any one month.

These considerations, and the fact that the selection was to be made by the terminating carrier, dictated the choice of waybill numbers "1" or ending in the digits "01" as the best selection device. This permitted the issuance of standard, simple, and unambiguous instructions which are so necessary for successful operation. The compounding device required to secure a one per cent selection from the small stations was a procedure established at the Commission for subsampling the "1" bills from these small stations with a probability proportional to the average number of bills issued per month over a recent 12-month period. This method is roughly equivalent to weighting the received bills; the expected values will be the same although the variances will be somewhat higher because of the discarded observations. The difference is relatively small and is more than compensated for by increased simplicity in subsequent operations.

The excess bills over the required one per cent, which result from monthly and annual numbering systems, are an unnecessary cost to both the carriers and the Commission and so the carriers have been urged to change to the more desirable block system. Many have done so and consequently the problems associated with establishment of a proper sample design for small stations are becoming less important.

After developing and applying a correct sample design, it is evident that a satisfactory sample depends upon the complete and accurate selection of the designated waybills. These waybills are selected by the reporting carriers in over one hundred offices through the occasional efforts of perhaps a thousand individuals. This factor introduces possibilities of error which require constant policing to assure the desired accuracy. Fortunately, the Commission receives another report on carload traffic which can be compared to the waybill sample as an effective check for completeness and accuracy.

This report is known as the quarterly freight commodity statistics and shows for each railroad the total number of cars terminated by that road for each of the 261 carload commodity classes divided between traffic which originated on the reporting road and traffic which was originated on other carriers and was delivered to the reporting road for termination. Each quarter the waybill sample is tabulated in

this same manner for each reporting carrier and the results, commodity by commodity, are compared to 1 per cent of that road's freight commodity statistics. Any significant discrepancies noted in this comparison are carefully reviewed and the causes for such discrepancies are determined. It has been interesting to note that some of the discrepancies have been due to errors in the commodity statistics rather than in the sample and so in these cases the sample has actually been used as a quality control for the 100 per cent report. An idea of the effectiveness of the control of the sample may be gained from the fact that in each quarter there are in the neighborhood of 20,000 individual comparisons. The sample itself includes about 75,000 observations.

The selection of waybills on the basis of the waybill number provides in effect for a sample of every hundredth bill issued by each station. This is actually true in the case of stations issuing bills on a block system and in the case of stations issuing many hundred bills per month and is approximated at the others. Consequently the sample is quite efficient because of the relationship between the originating station and important transportation factors such as the commodity and territory. The periodic comparison of the sample to control totals helps reduce selection errors and assures its validity. The sample has found many uses in varied fields and its application to one transportation problem will be discussed, but before doing so it might be well to consider the foregoing procedures in the light of modern sampling theory [1].

The universe considered here is the totality of rail carload shipments terminated by Class I railroads in the United States. The sampling unit is the waybill, which in general covers a single carload, and the frame is the complete series of numbered waybills. These numbers are assigned at the origin station so that when a block numbering system is used selection based on the terminating digits "01" will provide a systematic sample of every hundredth waybill issued, stratified by origin. Since the shipment itself is independent of the terminating digits of the waybill number this will result in an unbiased, random—and therefore representative—one per cent sample of carload shipments.

A modification of this design is necessary in the case of small stations numbering their bills on a monthly system. Where at least one but less than 100 bills are issued per month the selection of the "1" bills provide a sample consisting of the first shipment in the month for each station. This, therefore, is not necessarily a random selection except in the sense that there is some chance element in whether or not a particular ship-

ment will be the first one. The essential random element is provided by a further process of subsampling these initially selected bills.

The "1" bills from each of the small stations over several months form a series which approximates a systematic sample with a sampling interval equal to the average number of bills issued per month. The desired interval is 100 and this in turn can be approximated by discarding a sufficient number of bills so that the resulting average interval is 100. The bills to be discarded are selected by a chance process with appropriate probabilities so the final sample is a random one. This permits valid estimates of standard errors, an example of which is given later.

There is a problem of non-response due usually to occasional failure to select all the required "1" and "01" bills. These failures are essentially random in character and do not appreciably affect computed averages or ratios although of course they do introduce a negative bias in estimates of aggregates. The quarterly comparison to the 100 per cent count of the commodity statistics report provides an excellent control and good progress is being made in the establishment of accurate mechanical checks by the reporting carriers which reduce the non-response error.

Many diverse uses have been made of the Commission's waybill sample but it is proposed to examine the technical aspects of its use in only one problem. This problems, however, is one of sufficient complexity to indicate the value of the sample as a research tool.

The need for an index which would measure changes in rail carload freight revenue resulting solely from increases or decreases in freight rates has been recognized for years. Experimental indexes for various commodities were developed by the Commission but the methods used were too costly to permit any extensive expansion of these series. The Commission's 1 per cent waybill sample provided a new approach to the problem and offered the hope that annual indexes of many freight rates could be prepared and kept current with a relatively small expenditure of time and funds. The method finally adopted uses, with minor modifications, information contained in the regular releases of waybill statistics.

The regular Commission waybill statistics are tabulations of punch cards prepared from the waybill samples reported by the terminating carriers. The illustration on page 238 shows facsimiles of a typical waybill report and the detail and waybill cards prepared from that report. It also gives a brief description of the processing of the bills and cards.

Mileage block statistics have been developed by classifying the

traffic according to traffic categories determined by the commodity class, short-line length of haul, type of rate, and territorial movement. These are the statistics used in the preparation of the freight rate indexes. It should be noted that the rates, or prices charged for transportion, are related to each of these factors and they were chosen with a view towards making the items within each resulting traffic category as homogeneous as possible. The sample shipments for each year fall into about 30,000 of these traffic categories.

The indexes are constructed by comparing the tonnage and average revenue per ton of matching categories in the base and comparing year. Two sets of revenue figures are prepared from these data. The first consists of the actual revenue for the base year which, of course, is simply the total of the revenues for each of the comparing categories in this year. A second set of revenue figures are obtained by applying the average rate per ton for each category in the comparing year to the tonnage which moved in the same category for the base year. This produces a figure which represents the revenue which the traffic moving in the base year would have produced at the average rates in effect in the comparing year. The ratio of these two revenues yields the desired index figure.

It is evident that factors other than changes in freight rates might affect the average revenue per ton for a given traffic category. For example, the average length of haul within the mileage block assigned to the traffic category might be greater or less in the comparing year than in the base year. Similarly, the specific commodities which moved within the given commodity class might be different and take different rates. Therefore, the average revenue effect of such differences must be small as compared to the effect of rate changes if the resulting indexes are to be a satisfactory measure of changes in rates. Careful study of this problem does indicate that this requirement is substantially satisfield. As noted before, the traffic categories were chosen so as to be as nearly homogeneous with respect to rate characteristics as practicable. The milage blocks are shorter where changes in rate progression are most rapid. The various territorial movements are kept separate as are movements on interstate and intrastate rates and there is a further classification by commodity class and type of rate. Consequently, the area of fluctuation in average revenue from causes other than changes in the rate is relatively small in each of the traffic categories.

It is reasonable to expect that even when such fluctuations as these do occur, they are as likely to affect the results in one direction as another so that there is a tendency for the errors to compensate rather than to accumulate. In consequence, the net effect is that the changes in

average revenues noted by this method will usually reflect quite closely the revenue effects of actual changes in rates. Therefore, the indexes will be a measure of the average changes in rates. An illustrative page from the published statement is shown in Table 1.

These indexes are based upon only one of the 100 similar 1 per cent samples which could have been selected from the waybills representing the total carload traffic. Presumably indexes computed in exactly the same manner from each of the 100 different samples would have produced 100 slightly different results. Each of these in turn would have been an estimate of the index which could theoretically have been produced from a 100 per cent sample consisting of all of the waybills. This condition, of course, immediately raises the question of how much does a particular index computed from the sample which actually was selected differ from the theoretical index which could have been prepared from the total of all waybills.

There are a number of methods by which an estimate of the standard deviation of any computed index could be obtained but in particular there are two relatively simple procedures which were actually applied to the computed rate indexes and which have been used for similar determinations of the confidence which could be placed in other estimates prepared from the waybill sample.

It was noted that you would expect somewhat different results from computations made on each of the 100 possible 1 per cent samples which could have been drawn from the total of all waybills. If the variation in these estimates could be known, it would be easy to calculate the standard deviation for one of them. Obviously, it is not possible to get such figures as these but an extension of the reasoning can be made to the sample which is available. If this sample is divided up into, say, ten parts in a manner similar to the method used in the selection of the original sample, then the variation observed in separate estimates made from each of these ten parts can be used to estimate the possible variation in the estimate prepared from the whole sample [1]. If $X_1 \cdots X_{10}$ are the ten subsample estimates then the variance of the total sample estimate $\overline{X}$ is approximately:

$$S_{\bar{x}}^2 = \frac{1}{10}\,\frac{1}{9}\sum_{1}^{10}(X_i - \overline{X})^2$$

or, using the range $R = X_{max} - X_{min}$ the standard deviation can be estimated from the relationship [2]

$$S_{\bar{x}} = \frac{.325}{\sqrt{10}}R = .103R.$$

TABLE 1

INDEXES OF AVERAGE FREIGHT RATES FOR COMMODITY GROUPS AND SELECTED COMMODITY CLASSES*
(1950 = 100)

Item	Index						Per cent Increase 1952 Over 1947	Approx. Standard Deviation
	1947	1948	1949	1950	1951	1952		
All Commodities	80	93	99	100	102	109	36	.5
Group I—Products of Agriculture	80	93	98	100	102	108	35	.5
Class 001 Wheat	80	93	95	100	102	108	35	1.0
003 Corn	75	94	95	100	102	109	45	1.0
033 Cotton in Bales	82	92	98	100	102	112	37	1.0
† Oil Bearing Crops	78	91	99	100	103	111	42	4.0
† Fresh Fruits	83	95	99	100	101	106	28	.5
† Fresh Vegetables	82	95	99	100	101	106	29	.5
085 Potatoes, Other Than Sweet	79	94	99	100	101	107	35	1.0
101 Sugar Beets	86	96	103	100	106	110	28	3.0
199 Products of Agriculture, N.O.S.	88	94	105	100	101	116	32	3.0
Group II—Animals and Products	77	93	99	100	102	110	43	.5
Class 203 Cattle and Calves, S.D.	79	94	98	100	103	112	42	1.0
215 Meats, fresh, N.O.S.	73	90	98	100	103	110	51	1.0
Group III—Products of Mines	83	91	98	100	102	108	30	.5
Class 301 Anthracite Coal, N.O.S.	81	90	99	100	101	106	31	.5
305 Bituminous Coal	81	89	98	100	102	107	32	.5
307 Coke	83	91	97	100	103	108	30	1.5
309 Iron Ore	88	95	99	100	103	110	25	.5
323 Clay and Bentonite	84	95	98	100	104	111	32	1.0
325 Sand, Industrial	77	91	98	100	105	113	47	1.0
327 Gravel and Sand, N.O.S.	90	98	102	100	102	107	19	1.0
329 Stone and Rock: Broken, Ground and Crushed	85	93	100	100	103	108	27	1.0
331 Fluxing Stone & Raw Dolomite	75	88	98	100	104	109	45	1.0
337 Petroleum, Crude	81	96	99	100	102	109	35	1.5
339 Asphalt	77	94	100	100	101	110	43	1.0
341 Salt	86	93	98	100	102	108	26	1.0
343 Phosphate Rock	81	95	100	100	104	110	36	1.5
399 Products of Mines, N.O.S.	82	98	100	100	103	112	37	2.5
Group IV—Products of Forests	79	93	98	100	102	109	38	.5
Class 401 Logs, Butts, and Bolts	83	93	99	100	103	108	30	1.5
403 Posts, Poles, & Piling, Wooden	79	92	97	100	101	111	41	1.5
409 Pulpwood	88	98	100	100	104	110	25	.5
422 Lumber, Shingles, and Lath	78	92	98	100	102	109	49	.5
499 Products of Forests, N.O.S.	83	95	100	100	104	112	35	2.5

* Only classes having more than approximately 1,000 cars in the sample are shown separately. The group indexes cover all classes in the group.

† These groupings of individual commodity classes are shown at the request of the Department of Agriculture. "Oil bearing crops" include commodity classes 037 to 047, 097 and 105; "Fresh fruits", classes 049 through 069; "Fresh vegetables" classes 077 through 089 (including 085 which is also shown separately).

This latter estimate is only about 85 per cent as efficient as the former but has the great advantage of simplicity and ease of computation.

The first estimates of the standard deviation of the rate indexes were made by following this procedure in principle. It was recognized, however, that the preparation of ten separate sets of indexes from ten subsamples would be too great a task for the limited staff available. Consequently, the individual waybills in the sample were divided into only five groups. This division was made as nearly as possible on the same basis as that by which the initial sample was selected. Each of these five groups was classified according to traffic categories and five different sets of indexes were computed. The variation in these indexes then yielded estimates of the standard deviation for the indexes prepared from the total sample. This method, however, proved to be an even greater task than initially contemplated and a less accurate but substantially less costly procedure was developed.

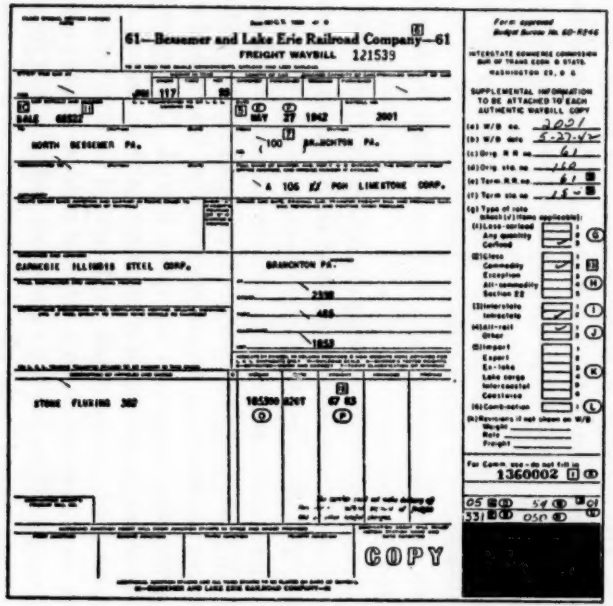
The second procedure made use of the same mileage block cards that were used in developing the initial indexes. These cards were randomly distributed into ten groups and the indexes prepared as before for each of the groups. The variation in these indexes then yielded an estimate of the standard deviation of the index prepared from the total of the mileage block cards. While this procedure is not as accurate as the one using the individual waybill cards, the error is on the safe side because there is a bias which tends to indicate a larger value of the standard deviation. However, the substantial savings resulting from the use of mileage block cards instead of the individual waybill cards more than offsets the loss in accuracy.

Incidentally, the mile block cards were randomly distributed into ten subgroups by means of a one column sort on the units position of the revenue field. The digits in this field are equivalent to the sum modulo ten of the units digits in the waybill cards included in each block. It has been shown [3] that such a process of "compound randomization" will yield sequences approaching equal probabilities for each digit, even when the generating sequences are quite biased. Consequently it is permissible to use aggregate fields of summary cards in this manner even when the same fields in the detail cards would be unsatisfactory. Extensive tests have been made on a table of random digits [4] produced in a similar manner from waybill data which indicate that considerable confidence can be placed in the effective randomness of the results.

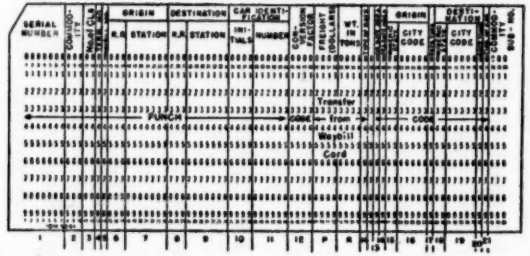
The results of the two procedures for estimating the index standard deviations were in substantial agreement although individual examples of wide variation did appear. These were usually due to unequal distributions of traffic in the mileage block cards which of course is one

INTERSTATE COMMERCE COMMISSION
BUREAU OF TRANSPORT ECONOMICS AND STATISTICS

CONTINUOUS CARLOAD WAYBILL STUDY
CARD FORMS AND PROCEDURE



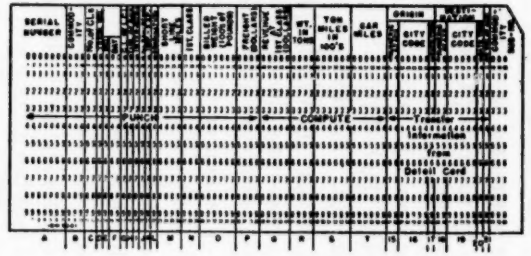
DETAIL CARD



Indicates punched in detail card

Indicates punched in waybill card

WAYBILL CARD



Immediately upon receipt the waybills are reviewed for completeness and legibility and the number "1" waybills issued by stations having monthly numbering systems are subsampled to eliminate the bias resulting from the numbering system. The short-line distance and first-class rate between origin and destination stations are then checked and the bill is coded for commodity, month of termination, and number of carloads. It is then serially numbered and two cards are punched as shown above.

The detail card is used for mechanically coding the origin and destination and type of car. It is first matched against a station master file which contains a city, state, and geographical area code which is reproduced into the detail card for both origin and destination. The geographical area codes are used to develop the territorial movement which takes into account the direction of movement. The detail card is then matched against the car initial and number master file which contains the type of car code and which completes the coding.

The waybill card is the basis for computing revenue at first-class, weight in tons, ton-miles and car-miles. When the-e multiplications have been completed and verified the two cards are matched against each other and information transferred from one to the other as indicated.

Tabulations can be prepared from either or both cards. Through the use of summary cards, additional computations and analyses can be made for special characteristics of traffic.

of the prices which must be paid for using this approximation method. However, as was expected, the discrepancies were usually on the side of overestimating the error and so the less costly procedure is considered a safe one for determining the confidence which can be placed in the estimates.

The preceding discussion has covered the selection and adjustment of the carload waybill sample, and the method used to check its completeness and accuracy by comparison with the carriers' freight commodity statistics report. It has shown how the regularly published waybill statistics can be used, with minor modifications, to get the answer to a complex transportation problem in a simple and straight forward manner. As a final example several methods for estimating the standard deviations of sample results were developed. In each case the techniques used were simple and direct. This is characteristic of the current waybill sample and illustrates its value as a tool for transportation analysis.

REFERENCES

[1] Deming, William Edwards, *Some Theory of Sampling*, New York, John Wiley & Sons, Inc., (1950).

[2] Dixon, Wilfrid J., and Massey, Frank J., Jr., *Introduction to Statistical Analysis*, New York, McGraw-Hill Book Co., Inc., (1951).

[3] Horton, H. B., and Smith, R. T., "A direct method for producing random digits," *Annals of Mathematical Statistics*, 20 (1949), 82–90.

[4] Horton, H. B., and Smith, R. T., *Table of 105,000 Random Decimal Digits*, Washington, D. C., Interstate Commerce Commission, (1949).

RESPONSE ERRORS IN THE COLLECTION OF WAGE STATISTICS BY MAIL QUESTIONNAIRE

SAMUEL E. COHEN AND BENJAMIN LIPSTEIN
U. S. Bureau of Labor Statistics

THE literature on response errors in mail questionnaires has for the most part been devoted to the personal characteristics of the individual, e.g. age, education, attitude toward social questions, etc. Little attention appears to have been given to response errors where the data are collected from firms or organizations. It is doubtful, however, that a sharp separation can be made between response errors of personal characteristics and data on business operations. In the small firm, elements of prestige may exist with respect to exaggerating sales, production, or employment. The large firm on the other hand may be reluctant to disclose the extent of its operations for competitive reasons, or in government surveys may fear that the information may be used for other than statistical purposes. Further, during periods of economic control, some business firms may consider it advantageous to bias reports in their own favor for future administrative action. Thus, it appears that a parallel may exist between response biases of individuals and the response bias of firms.

The problem errors of response of firms to a mail questionnaire was recently studied in a wage survey of the rainwear industry conducted by the U. S. Bureau of Labor Statistics. Two kinds of response errors were examined, 1) coverage errors and 2) errors in the reporting of wage data. Coverage error as part of response error exists in this instance because the firm classifies itself as being in or outside of the industry. It includes firms that were incorrectly classified as being in the industry as well as those which were incorrectly excluded.

Errors in reporting of wage data include such errors as incorrect reporting of wage rates, inclusion of workers outside the scope of survey (e.g.—office workers), and erroneous division into the auxiliary[1] or non-auxiliary class.

The survey was initially made during July 1951 for a June 1951 payroll period, to provide data for aiding in the determination of the pre-

[1] For the purpose of this survey auxiliary workers are defined as those workers engaged in one or more of following occupations: position marking, shade and size numbering, bundle tying, bundle ticketing, matching and pairing, basting, pulling, hand trimming, cleaning, turning; floor boys and girls, porters, examiner's helpers. Non-auxiliary workers consist primarily of cutters and sewing machine operators.

vailing minimum wage in that industry.[2] Responses were obtained predominantly by mail questionnaire. A subsample of the non-respondents to the mail questionnaire was contacted in person or by telephone so that an unbiased estimate of the wage distribution of the industry could be prepared. Subsequently, questions were raised as to whether mail collection was appropriate, and whether the data presented were acceptable for use in minimum wage determination. Accordingly, a field resurvey was made to check the survey results and to determine whether mail questionnaires are feasible in surveys of such industries as rainwear, where piece-work payments prevail.

The original respondents were given no advance warning of the resurvey. Difficulties generally encountered in studies of response error because of uncertainty in identifying the reporting unit were not present in this survey. It was possible to identify the reporting units as being identical in both surveys—something not always possible with families or households.

The variables measured in both the survey and resurvey were subject to objective determination from payroll records. Thus variations in reporting would be attributed to indifference, misinterpretation of instructions, arithmetic errors or purposeful concealment. These items cannot be isolated in most studies of survey techniques and reporting practices. For example, differences in reporting educational attainment can be attributed in part to the uncertainty of the definition. A very common source of variation in reporting, age, is generally not subject to any definite verification. All that is available as a rule are two replies by the same respondent, both of which may be in error.

SURVEY METHODS

The universe of establishments in the rainwear industry was defined to include all establishments whose value of product of rainwear was 50 per cent or more of their total production in 1950 or who maintain separate rainwear departments. An industry of this type is not easily defined or isolated from the remainder of the needle trade industries. A comparatively small number of firms are continuously in the industry from year to year. For the most part, firms shift in and out of this industry with the fluctuations in product demand. The plant

[2] The Walsh-Healey Act specifies that employers engaged in government contracts exceeding $10,000 must pay the prevailing minimum rate of pay for such work. The Secretary of Labor is assigned the responsibility of determining the prevailing rate of pay in question. Administratively, this work is carried out by the Wage and Hour and Public Contracts Division of the Department of Labor. The Bureau of Labor Statistics is frequently called upon to conduct surveys of wage rates in the industry. The data collected in these surveys are used as an aid in this wage determination.

equipment of firms producing rainwear is not specialized, but rather is equally usable for many other needle trade products.

The initial listing of firms comprising the industry was obtained from Unemployment Compensation listings, supplemented by lists of firms provided by the International Ladies' Garment Workers' Union and the Amalgamated Clothing Workers Union.[3] The first wave of questionnaires was mailed to all listed firms (633). A second wave of schedules was sent to all nonrespondents. The 353 responses to these two mailings consisted of 84 firms being defined as within scope (firms within the rainwear industry) and 269 firms as being out of scope. Of the nonrespondent group, a sample of 70 firms was selected for direct contact by telephone or personal visit. The non-response sample produced 19 firms within the scope of the survey and 51 out of scope.

TABLE 1

SAMPLING PROCEDURE OF ORIGINAL SURVEY

Stages in the Sampling Procedure	Number of Schedules	In Scope	Out of Scope
Total mailed	633		
Responded	353	84*	269
Did not respond	280		
Direct contact	70	19†	51
Total received	423	103	320

* 66 schedules usable—rejected schedules failed to provide wage distributions or other essential data.

† All schedules usable.

The resurvey consisted of two parts, 1) a sample check of firms reported initially as out of scope, and 2) the collection by personal visit of data initially obtained by mail.

Part (1) of this investigation consisted of a sample of 42 firms in New York City classified as out of scope on the basis of mail returns in the initial survey. These 42 firms were rechecked by personal visit. Of 17 firms initially reported as out of business, of the sample of 42, it was found that 10 were completely out of business, one was manufacturing another product, 4 were defunct rainwear subsidiaries of firms still in other lines, and 2 were in the rainwear business elsewhere and under different names (included in the universe under the new name).

[3] Past experience has indicated that lists compiled in this manner contain most of the firms in the industry. An exhaustive investigation of the completeness of the list would be financially prohibitive, as it would involve a detailed investigation of a number of related industries.

Of the 25 remaining firms in the out of scope sample—only 3 could be classified properly in the rainwear industry as defined.

On the basis of this sample survey, it appears that the Bureau underestimated the number of rainwear establishments in New York City by about 15 or approximately 20 per cent.

Most of these 42 establishments at some previous time manufactured rainwear. However, a general decline in the volume of such production had resulted in rainwear no longer being the principal product or the

TABLE 2

RESULTS OF RESURVEY OF FIRMS CLASSIFIED AS OUT OF SCOPE IN ORIGINAL SURVEY, NEW YORK CITY

Classification	Number of Firms	
	Original Survey	Resurvey
Total Sample of Firms	42	42
Out of Business	17	10
Change of product		1
Discontinued rainwear dept.		4
Operating, under different names:		
Included in universe		2
Manufacturing other products	25	22
Manufacturing rainwear		3
Total out of scope	42	39

discontinuance of a separate rainwear department. Such shifts had been taking place since 1946 and explain in part the divergence of the survey data with respect to number of establishments (1951) from that of the 1947 Census of Manufactures.

Despite the under-representation of the Middle Atlantic region (as suggested by the New York City results) in the Bureau of Labor Statistics survey, no great difference in the Nation-wide statistics could be generated by such a deficiency. Under the extreme assumption—that the bias of omission occurs in that area alone—the increase of the Middle Atlantic employment raises the Nation-wide median by 1 cent and decreases the percentages of workers in the lowest class intervals (75 cents–80 cents) by no more than three-tenths of 1 per cent. Actually, an understatement of 20 per cent in number of firms in New York City

and in turn in the Middle Atlantic Region means less than a 20 per cent understatement in employment since firms in this region are smaller than those in other regions. Such field work as was done in New York is very expensive to do elsewhere because of the scatter of the remaining establishments. A small sample in other areas would be practically useless because of the apparently low proportion of bona fide rainwear firms erroneously excluded.

Do errors of the universe, such as we have just described, form an effective argument against making mail surveys in such industries as rainwear, where the establishment turn-over is so great? Without the expenditure of large sums, it is likely that the error in any field survey would be even greater. In order to obtain 85 usable schedules, it was necessary to make an initial mailing to more than 600 establishments, most of which did not fall within the scope of the survey. The field experience would not be appreciably better than this, judging from the canvass made in New York. Such a canvass by personal visit would be considerably more expensive. Hence the alternatives presented are (1) a mail survey adjusted for non-response, or (2) no survey at all. The extent to which correction should be made depends on the use to be made of the survey.

In the second part of the resurvey, an attempt was made to visit all firms not originally interviewed in person. However, visits were not made by the Bureau to 13 establishments which were included in a field survey made by Dr. Lazare Teper of the International Ladies' Garment Workers' Union. Dr. Teper provided the Bureau with data collected in his survey.

Wage data collected in the resurvey were obtained directly from payroll records. Schedules were obtained from 53 of the 57 firms visited by by Bureau of Labor Statistics. One firm employing about 100 employees refused to cooperate in the resurvey. The other three, employing a total of 28 workers, were either out of business at the time of the resurvey, or their records were unavailable.

The retabulations are thus based on 15 schedules obtained originally by personal visit, 13 obtained by the ILGWU, and 53 obtained in the resurvey.[4]

All schedules were given the same weight in the retabulation as in the original survey. The final tabulations are therefore the equivalent of a complete field survey.

---

[4] Four of the original 19 schedules obtained by direct contact as mentioned in Table 1 were obtained by telephone. These were included in the 53 firms resurveyed.

COMPARISON OF RESULTS

In comparing the results of the two surveys, it is assumed throughout that the data collected by personal interview are without errors. In reality, it is never possible to obtain absolutely accurate data whether by field or by mail procedures. Errors and biases exist in all data in varying degrees. Considering the nature of the survey, it is reasonable to assume that the data collected by personal interview are freer from biases of reporting and random errors than those obtained by mail.

Table 3 and Chart 1 show the comparative cumulative distribution of workers by earnings obtained by the two methods of collection.

TABLE 3

CUMULATIVE PERCENTAGE DISTRIBUTION OF PRODUCTION WORKERS IN THE RAINWEAR INDUSTRY BY STRAIGHT-TIME AVERAGE HOURLY EARNINGS, UNITED STATES AND SELECTED REGIONS, JUNE 1951, ORIGINAL AND RESURVEY

Average Hourly earnings* (in cents)	United States†		New England	
	Original	Resurvey	Original	Resurvey
Under 75.0	‡	0.4	—	0.7
Under 80.0	13.4	11.8	17.3	18.5
Under 85.0	17.2	16.2	22.4	23.4
Under 90.0	23.7	22.6	27.8	28.4
Under 95.0	28.6	27.7	32.3	32.9
Under 100.0	34.0	34.1	35.4	36.3
Under 105.0	42.2	41.4	43.5	43.6
Under 110.0	48.1	47.0	48.0	48.2
Under 115.0	55.0	54.1	54.0	53.1
Under 120.0	59.2	58.5	58.0	57.6
Under 125.0	63.6	63.3	61.6	61.6
Under 130.0	68.7	67.8	66.1	65.6
Under 135.0	72.4	71.6	68.9	68.6
Under 140.0	74.7	74.1	71.0	70.9
Under 145.0	77.6	77.1	74.3	74.0
Under 150.0	79.8	79.0	76.5	76.1
Under 155.0	82.4	81.4	79.9	79.1
Under 160.0	83.8	82.9	82.0	81.5
Under 165.0	85.6	84.8	84.3	83.9
Under 170.0	86.7	86.0	85.6	85.4
Under 175.0	87.8	87.3	87.0	87.0
Number of workers	9,149	8,929	4,005	3,960
Median rate	$1.10	$1.12	$1.12	$1.12

TABLE 3—(*Continued*)

Average hourly earnings* (in cents)	Middle Atlantic		Great Lakes	
	Original	Resurvey	Original	Resurvey
Under 75.0	‡	0.3	0.1	0.1
Under 80.0	9.2	4.8	10.8	7.1
Under 85.0	12.4	9.0	11.4	9.0
Under 90.0	17.0	13.9	21.5	19.5
Under 95.0	21.5	19.3	27.2	25.4
Under 100.0	24.8	24.0	38.5	38.6
Under 105.0	34.5	32.1	45.6	45.1
Under 110.0	39.0	37.0	55.1	53.4
Under 115.0	43.5	42.2	65.2	65.7
Under 120.0	45.4	44.0	72.2	72.6
Under 125.0	49.0	48.0	79.0	79.4
Under 130.0	55.5	53.1	83.9	84.1
Under 135.0	60.7	58.6	87.7	87.8
Under 140.0	63.6	61.6	89.8	89.9
Under 145.0	66.2	64.8	92.6	92.8
Under 150.0	68.8	66.3	94.8	94.6
Under 155.0	71.5	68.8	95.7	95.6
Under 160.0	72.3	69.9	96.3	96.3
Under 165.0	73.9	71.5	97.4	97.4
Under 170.0	75.6	73.3	97.8	97.7
Under 175.0	76.7	74.8	98.4	98.4
Number of workers	2,276	2,205	2,527	2,432
Median rate	$1.27	$1.27	$1.07	$1.08

* Excludes premium pay for overtime and night work.
† Includes data for other regions in addition to those shown separately.
‡ Less than .05 of 1 per cent.

Beyond the 90-cent point there is little difference in any of the regional data, and even below this point there is little difference for the industry as a whole. There were slight increases in the median rates, 2 cents in the industry as a whole, 2 cents in the Middle Atlantic region, and 1 cent in the Great Lakes region.

The lower end of a wage distribution is of critical importance for minimum wage determination. Examination of these parts of the distribution shows that the greatest difference occurs in the interval between 75 cents and 80 cents, 11.4 per cent of the workers were classified in this interval in the resurvey as against 13.4 per cent in the original survey. At no other point does the difference exceed 1 percentage point. Inspection of the cumulative distribution (Table 3 and Chart 1) show very small differences beyond the 90-cent point.

The greatest discrepancy between the surveys occurs in two of the

regional distributions. In the Middle Atlantic region, the estimated percentages of workers between 85 cents and 90 cents were 9.2 per cent for the original survey and 4.5 per cent in the resurvey. In the Great

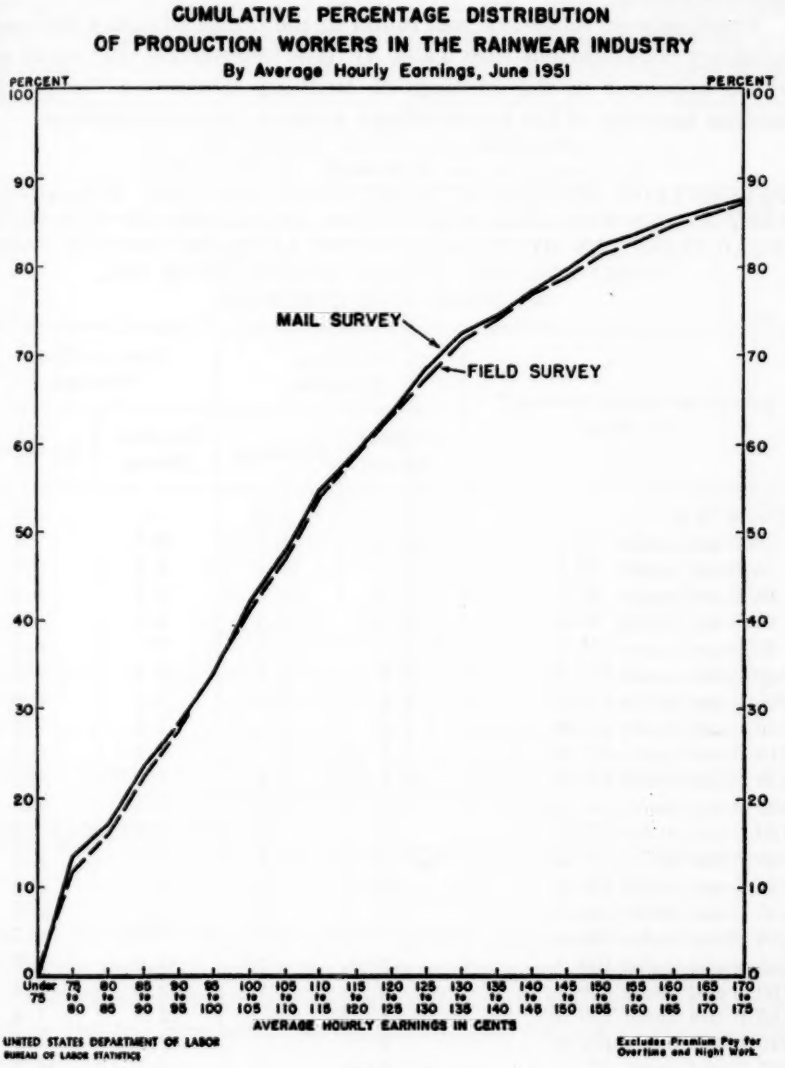

CUMULATIVE PERCENTAGE DISTRIBUTION
OF PRODUCTION WORKERS IN THE RAINWEAR INDUSTRY
By Average Hourly Earnings, June 1951
PERCENT
PERCENT
MAIL SURVEY
FIELD SURVEY
AVERAGE HOURLY EARNINGS IN CENTS
UNITED STATES DEPARTMENT OF LABOR
BUREAU OF LABOR STATISTICS
Excludes Premium Pay for Overtime and Night Work.

CHART 1

Lakes region the corresponding percentages were 10.7 and 7.0. These differences are due primarily to the two firms reporting large blocks of workers at the guaranteed rate rather than at actual earnings, plus

two other questionable schedules with concentrations of workers at 75 cents in the original survey. (Of these, one firm refused to cooperate in the resurvey, while the other firm supplied incorrect data in the original survey.)

Practically no difference was found in the broad auxiliary and non-auxiliary classifications (See Table 4). It is noteworthy that small declines appeared in the earnings of auxiliary workers, although the median earnings of the non-auxiliary workers increased slightly.

TABLE 4

PERCENTAGE DISTRIBUTION OF AUXILIARY AND NON-AUXILIARY PLANT WORKERS (EXCLUDING LEARNERS) IN THE RAINWEAR INDUSTRY BY STRAIGHT-TIME AVERAGE HOURLY EARNINGS,* FOR THE UNITED STATES, JUNE 1951, ORIGINAL AND RESURVEY

Average hourly earnings* (in cents)	Auxiliary Workers		Non-Auxiliary Workers	
	Original Survey	Resurvey	Original Survey	Resurvey
Under 75.0	0.1	0.3	—	0.4
75.0 and under 80.0	36.3	41.8	9.6	7.7
80.0 and under 85.0	10.5	10.7	3.1	3.6
85.0 and under 90.0	11.9	14.7	5.8	5.4
90.0 and under 95.0	5.8	7.5	4.8	4.8
95.0 and under 100.0	4.1	5.1	5.6	6.5
100.0 and under 105.0	8.9	6.3	8.1	7.4
105.0 and under 110.0	3.6	2.4	6.1	6.0
110.0 and under 115.0	3.0	1.4	7.3	7.8
115.0 and under 120.0	2.7	3.9	4.4	4.5
120.0 and under 125.0	2.1	2.5	4.8	5.0
125.0 and under 130.0	2.0	1.2	5.5	4.9
130.0 and under 135.0	1.3	.8	4.2	4.2
135.0 and under 140.0	1.1	.4	2.6	2.7
140.0 and under 145.0	.7	.3	3.3	3.4
145.0 and under 150.0	3.2	.3	2.3	2.1
150.0 and under 155.0	1.0	—	2.8	2.7
155.0 and under 160.0	.3	.2	1.5	1.7
160.0 and under 165.0	.5	.1	2.0	2.1
165.0 and under 170.0	.1	—	1.3	1.4
170.0 and under 175.0	.2	—	1.2	1.5
175.0 and over	.6	.1	13.7	14.2
Total	100.0	100.0	100.0	100.0
Number of workers	1,027	990	8,051	7,939
Median rate	$0.85	$0.84	$1.14	$1.15

* Excludes premium pay for overtime and night work.

REPORTING ERRORS OF FIRMS IN SCOPE

Establishments revisited gave a variety of explanations as to how the original data were compiled. Of the 53 firms studied, 25 followed the instructions completely and correctly; a few included salesmen and office workers; some estimated earnings; a few included overtime and 10 simply could not recall what was done. The most serious downward

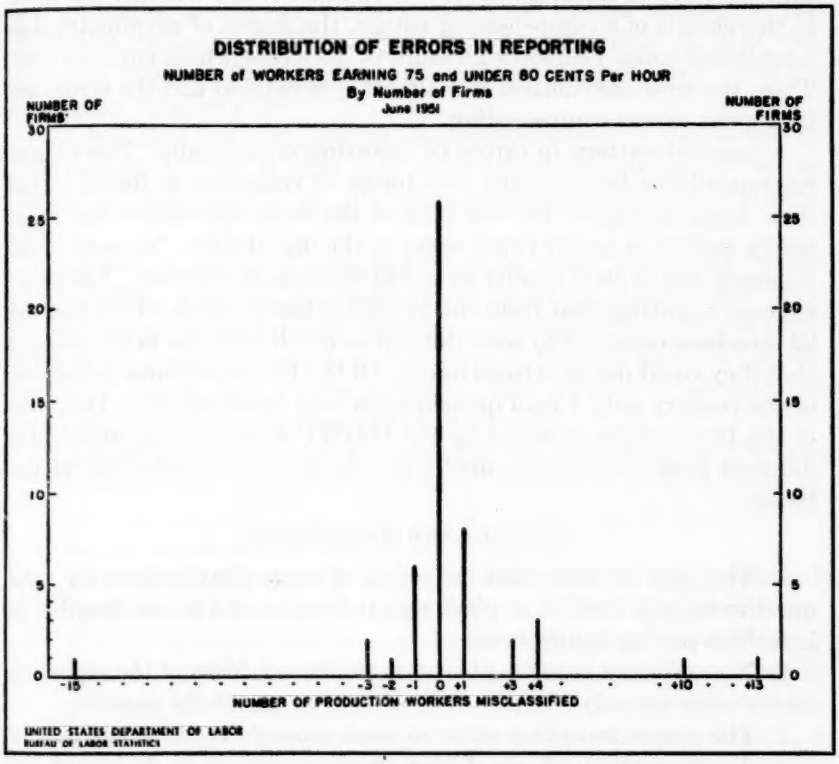


CHART 2

bias resulted from two firms which reported guaranteed rates instead of actual earned rates. The main difference between the original survey and the resurvey is due to the errors of these two firms and to the omission of the one firm with 100 employees that refused to cooperate in the resurvey.

A comparison was made of changes in reporting in the 75 and under 80 cent groups (omitting the two firms which misinterpreted the instructions.)

A frequency distribution of the errors was constructed (Chart 2). The X variable was the number of production workers misclassified in

the 75 and under 80 cent class. Positive errors were workers erroneously included in the class, and negative errors were those who should have been included. The Y variable was the number of firms which misclassified workers.

This frequency distribution of reporting errors showed an arithmetic mean of -0.4. Thus, the critical class was less than half a worker per firm too large in the original survey. In order to see whether the error in this class is of a compensating nature, the degree of asymmetry was established using Pearson's measure of skewness which equaled 0.05. Thus, the error distribution was almost symmetrical and the errors are to a great extent compensating.

A regional pattern in errors of reporting is noticeable. The closest correspondence between the two forms of collection is found in the New England region. For the bulk of the firms agreement was practically perfect at nearly every point in the distribution. No mail questionnaire was of such quality as would necessitate rejection. The poorest mail reporting came from the Middle Atlantic region where 5 of the 22 schedules collected by mail differed so much from the field resurvey that they could not be termed usable. Of the 10 comparisons in the rest of the country only 1 mail questionnaire was unsatisfactory. Only one of the 13 schedules obtained by the ILGWU resurvey was sufficiently different from the original mail schedule to warrant rejection of the latter.

SUMMARY AND CONCLUSIONS

1. The data indicate that collection of wage distributions by mail questionnaire is feasible in piece rate industries and is also feasible in industries paying hourly rates.

2. No consistent reporting bias was evidenced. Most of the reporting errors were not only compensatory, but were generally small.

3. The major reporting error in such surveys is the confusing of earned and guaranteed rates. Care must be taken in design of the questionnaire and in the accompanying instructions to distinguish between those two items. Careful editing will frequently disclose whether large concentrations of workers at any one rate are caused by reporting guaranteed rates instead of earnings. Personal visit may be necessary to clarify such cases.

4. A field follow-up of respondents classifying themselves as out of scope is necessary in an industry characterized by frequent changes in product. Without such a field follow-up an accurate estimate of the universe is difficult, and may result in overrepresentation of some segments of the industry in the survey totals.

INDUSTRIAL CLASSES IN THE UNITED STATES 1870 TO 1950

TILLMAN M. SOGGE
St. Olaf College

IN AN earlier article the writer classified the gainful workers in the United States for 1870 to 1930, and the labor force in 1940 into industrial classes.[1] This paper brings this classification up to date by adding the data for 1950.

The fact that the occupational classifications used in the Census of 1950 and the Census of 1940 were quite comparable made the task of adding the data for 1950 to the previous compilations comparatively easy.[2]

Some interesting developments have taken place during the decade of the 1940's, a period during which our economy was affected very markedly by war and postwar changes. In the main these changes have not altered the direction of trends indicated in earlier decades but the degree of change has in some instances been increased. Some of the more important changes from 1940 to 1950 are:

1. The number of persons for whom occupations were reported increased from 50,737,284 in 1940 to 57,632,879 or an increase of 13.5 per cent which is a considerably greater rate of increase than during the depression decade of the thirties.
2. The number of persons engaged in farming occupations continued to decline in spite of the large increase in the total labor force. Table I indicates that the number of farm laborers has declined from 6,143,998 in 1910 to 2,497,637 in 1950 or a decline of over 59 per cent in 40 years. The decline in relative importance of agricultural occupations is clearly evidenced by a decrease from 17.4 per cent of the total in 1940 to 11.8 per cent of the total in 1950.
3. Both proprietors and officials and professional classes have shown a significant increase in numbers during the past decade, but the greatest increase in numbers is indicated in the lower salaried

[1] Tillman M. Sogge, "Industrial Classes in the United States in 1940," *Journal of the American Statistical Association*, 39 (1944), 516-18. Parallel data for the years 1870 through 1920 were first presented by Professor Alvin H. Hansen in the two following articles (1) "Industrial Class Alignments in the United States," *Journal of the American Statistical Association*, 17 (1920), 417-25, and (2) "Industrial Classes in the United States in 1920," *Journal of the American Statistical Association*, 18 (1922), 503-06. The data for 1930 were first presented in an article by the writer entitled, "Industrial Classes in the United States in 1930," *Journal of the American Statistical Association*, 28 (1933), 199-203.

[2] Coverage of the different classes has been described in the earlier articles in this series listed in the footnote above.

TABLE I

INDUSTRIAL CLASSES

	1870	1880	1890	1900	1910	1920	1930	1940	1950*
Farm laborers	2,885,996	3,323,876	3,004,061	4,410,877	6,143,998	4,178,637	4,392,764	3,505,275	2,497,637
Farmers	3,000,229	4,282,074	5,370,181	5,770,738	6,229,161	6,463,708	6,079,234	5,238,049	4,337,719
Proprietors and officials	581,378	807,049	1,347,329	1,811,715	2,879,023	3,168,418	4,270,856	4,197,207	5,046,577
Professional	414,708	666,338	1,114,507	1,565,686	2,074,792	2,760,190	3,845,559	4,454,381	5,968,173
Lower salaried	309,413	529,473	965,852	1,329,928	2,393,620	3,985,306	7,116,814	8,071,201	11,253,501
Servants	975,734	1,075,655	1,454,791	1,453,677	1,572,225	1,270,946	1,999,133	2,840,021	1,414,732
Industrial wage earners	3,328,351	5,286,829	7,360,442	10,263,569	14,556,979	17,648,072	18,512,640	19,957,220	24,761,214
Unclassified	1,010,114	1,420,795	2,118,498	2,467,043	2,317,538	2,138,971	2,612,920	2,383,930	2,353,326
Total	12,505,923	17,392,099	22,735,661	29,073,233	38,167,336	41,614,248	48,829,920	50,737,284	57,632,879

TABLE II

INDUSTRIAL CLASSES

(Per cent)

	1870	1880	1890	1900	1910	1920	1630	1940	1950*
Farm laborers	23.1	19.1	13.2	15.2	16.1	10.0	9.0	6.9	4.3
Farmers	24.0	24.6	23.6	19.8	16.3	15.5	12.4	10.5	7.5
Proprietors and officials	4.6	4.6	5.9	6.2	7.5	7.6	8.7	8.3	8.8
Professional	3.3	3.8	4.9	5.4	5.4	6.6	7.9	8.8	10.4
Lower salaried	2.5	3.0	4.3	4.6	6.3	9.6	14.6	15.9	19.5
Servants	7.8	6.2	6.4	5.0	4.1	3.1	4.1	5.6	2.5
Industrial wage earners	26.6	30.4	32.4	35.3	38.2	42.4	37.9	39.3	42.9
Unclassified	8.1	8.2	9.3	8.5	6.0	5.1	5.4	4.7	4.1

* These figures do not include 1,366,064 experienced persons in the labor force for whom occupations were not reported. The data were compiled from the 1950 Census of Population, Volume II Characteristics of the Population: Part I, United States Summary, Table 124, pp. 1-261 to 1-266.

group and in industrial wage earners. The increase of nearly five million industrial wage earners during the past decade gives evidence of the expansion in our national economy which has taken place. Even though there has been this large increase in the number of industrial wage earners during the past decade, and an increase in relative importance of this group from 39.3 to 42.9 per cent of the total, it is significant to note that the importance of the industrial wage earners in relation to all the non-agricultural classes combined was less in 1950 than in 1920.

4. The number of servants in 1950 was less than half the number in 1940 dropping from 2,840,021 to 1,414,732. Percentagewise, servants constituted 5.6 per cent of the total in 1940 and only 2.5 per cent of the total in 1950.

STATISTICAL METHODS FOR POISSON PROCESSES AND EXPONENTIAL POPULATIONS

ALLAN BIRNBAUM*
Columbia University

1. INTRODUCTION

THE research worker and statistician frequently deal with phenomena in which events of some type occur randomly in time, or in which particles are randomly distributed in space. The Poisson process is the formal model of such phenomena. Furthermore, many phenomena which may naturally be represented by use of the exponential distribution or the Poisson distribution can alternatively be represented as Poisson processes and dealt with advantageously in this form. Statistical methods for the study of such phenomena can be as flexible and yet simple as the Poisson process model itself. A number of such statistical methods are described and illustrated below. Some of these methods were developed recently, while others are well-known but are described here briefly for comparison and completeness.

2. THE POISSON PROCESS

The results of any experiment in which observation is performed continuously and "events" (i.e., occurrences of any specified kind) are tallied, can always be described by a function $x = x(t)$, which gives the number of events observed, x, during the first t units of observation, for all values of t from 0 through T, the total amount of observation performed. Such an experiment, yielding an observed function $x(t)$, is a Poisson process if the events occur randomly in the sense of the following natural definition: given that any number x of events are observed in any amount t of observation, the points of occurrence of the x events are randomly (i.e., independently uniformly) distributed between 0 and t. Examples in which the Poisson process is a very accurate and useful model are the following:

Example 1. When a Geiger counter is used in an essentially constant environment, the counter tallies the number of times it is hit by radioactive particles. Here the amount of observation performed is measured in units of time during which the counter is operated. The Poisson process represents this phenomenon quite accurately. (The assumptions for the Poisson process will be violated seriously under extreme conditions: In an environment of very high radioactivity, the

* Research supported by the Office of Naval Research.

counter's inability to record a hit extremely soon after an earlier hit will seriously violate the independence assumption. If the environment contains radioactive materials with extremely rapid rates of decay, the assumptions of equal probabilities will be seriously violated.)

Example 2. If any homogeneous or well mixed material containing practically infinitesimal foreign particles or flaws is inspected, the amount of observation will be measured by the volume (or area or length) of material inspected, and the "events" will be the particles or flaws observed. One case is the use of a haemocytometer to count the number of blood cells distributed over an area on a slide. (For studies of the magnitude of deviations from the independence assumption due to "crowding" of cells, see [1] and the references given there.) Another case is that of inspection of manufactured materials by length or area for defects, if the manufacturing process is such that there is no appreciable dependence in the locations of defects and the average defect rate does not change appreciably during the period of inspection. Other applications, and some statistical methods discussed below, will be found in [18].

A Poisson process can be characterized in the following two simple alternative and equivalent ways:

(a) The "waiting times" u between successive events are independently distributed with the exponential density function

$$g(u) = (1/\theta)e^{-u/\theta} \qquad \text{for } u \geqq 0.$$

Here θ is the mean of waiting times:

$$E(u) = \theta.$$

(b) The increment $y = x(t_2) - x(t_1)$ of $x(t)$ on any interval of length $d = t_2 - t_1$ has the Poisson distribution

$$p(y) = e^{-d\lambda}\frac{(d\lambda)^y}{y!}, \qquad y = 0, 1, \cdots,$$

and the increments of $x(t)$ on non-overlapping intervals are independent. Here $d\lambda$ is the mean increment on an interval of length d. Thus λ is the mean rate of occurrences, and $\lambda = 1/\theta$.

All statistical questions concerning Poisson processes involve inferences about the value of the single parameter λ or θ of one process, or about the values of the respective parameters of several processes. (Methods for determining whether a given process is Poisson will not be discussed here; Davis [6] describes and applies several methods.)

The flexibility and simplicity of statistical methods applicable to

Poisson processes seems to stem from two sources. One of these is the simple way in which information from different experiments or observations on a single process can be combined. As an illustration, consider the life-testing procedures developed by Epstein and Sobel [9, 10]. It is assumed here that the lengths of life of a type of electron tube, say, are exponentially distributed with unknown mean life θ. (For some empirical tests of this assumption, see [6].) If one tube is observed until failure, then replaced by a new tube which is again replaced upon its failure, and so on, we have a sequence of observed lengths of life u from an exponential distribution. Hence the number of failures observed by any time t is a Poisson process $x(t)$. Now suppose that we place any number of such tubes under observation, and add new tubes or remove tubes from observation in a quite arbitrary manner. The information given by such an experiment is fully described as a single Poisson process, where t is now taken to be the number of tube-hours of life observed at any point in the procedure, and $x(t)$ is the number of failures observed up to that point. The feature illustrated here is called the additivity property of the Poisson process.

The second source of flexibility of statistical methods is the fact that we can shift back and forth freely between use of the two simple distributions which characterize the Poisson process, the exponential and the Poisson distributions.

3. ESTIMATION OF λ OR θ

As a basic example of the flexibility possible in the statistical treatment of data from such procedures, consider the problem of estimating the parameter λ of a Poisson process by means of a confidence interval.

Method 1. If a fixed amount t of observation is performed, then the number x of events observed has the Poisson distribution with mean $\mu=\lambda t$,

$$\Pr(x) = e^{-\mu}\frac{\mu^x}{x!}, \qquad \text{for } x = 0, 1, 2, \cdots.$$

By standard methods we can construct a confidence interval for μ, at a desired confidence level. An exact construction can be based on tables of the Poisson distribution, and extensive tables of such confidence intervals as well as tables of the Poisson distribution are available in [13, 17, 20].

Method 2. If observation of a Poisson process is continued until a specified number m of events has been observed, and if T is the amount of observation actually required to observe m events, then $2\lambda T$ has the

chi-square distribution with $2m$ degrees of freedom. A confidence interval for λ at a desired confidence level can then be constructed in the standard way, as described below.

It will be useful here and also in connection with other statistical problems discussed in the following sections to note that statistical problems dealing with variances of normal populations have direct analogues in problems dealing with parameters of Poisson processes. The analogue of the present problem is that of constructing a confidence interval for the variance σ^2 of a normal population given S^2, the sum of squared deviations from the known or estimated mean, with $2m$ degrees of freedom; the confidence interval is constructed by using the fact that S^2/σ^2 has the chi-square distribution with $2m$ degrees of freedom. Thus $2T/\theta$ and S^2/σ^2 have the same distributions, and so do the statistics $2T$ and S^2 if the corresponding parameters θ and σ^2 are equal. Here and in later problems we can take advantage of this correspondence to apply statistical methods developed originally primarily for a problem dealing with normal variances to a corresponding problem dealing with Poisson processes, and vice-versa.

To construct the desired confidence interval for λ, we select from tables of the chi-square distribution with $2m$ degrees of freedom, values C and D such that the probabilities of values less than C, and greater than D, are each $\alpha/2$. Then we can write

$$1 - \alpha = \text{Prob}\,\{C \leqq 2\lambda T \leqq D\} = \text{Prob}\left\{\frac{C}{2T} \leqq \lambda \leqq \frac{D}{2T}\right\}.$$

Thus $(C/2T, D/2T)$ is a confidence interval for λ, with confidence coefficient $1-\alpha$. The small bias of such confidence intervals could be removed at the price of complicating the procedure.

The following method of constructing point estimates of λ based on Method 2 was developed in [15]. We have with probability $1-\alpha$ that

$$C/2T < \lambda < D/2T.$$

Given this inequality, the maximum percentage deviation of λ from an estimate λ' is minimized by taking

$$\lambda' = \frac{C + D}{4T}.$$

This maximum percentage deviation is

$$\tau = \frac{D - C}{C + D}\,100\%.$$

τ is a quantity which decreases as the number m of events observed increases. Hence we can use tables of the chi-square distribution to determine the number m of events to be observed to give an estimate λ' with maximum percentage deviation not exceeding any given positive τ, at confidence level $1-\alpha$.

On the other hand, when Method 1 is used, there is no way of prescribing the amount t of observation so as to obtain point estimates of bounded percentage error. As shown in most statistics texts, an approximate construction of the confidence intervals of Method 1 is given, say at the .95 confidence level, by the interval

$$\frac{x}{t} - \frac{1.96\sqrt{x}}{t}, \qquad \frac{x}{t} + \frac{1.96\sqrt{x}}{t}.$$

It is seen that here the maximum percentage deviation of λ from the estimate x/t is $(1.96/\sqrt{x})$ 100%, a function of the observed value of x.

Method 3. In Method 1, the absolute deviation of the estimate x/t from λ is bounded by $C\alpha(\sqrt{x}/t)$, where $C\alpha$ is a constant corresponding to the confidence level used, with probability $1-\alpha$. In Method 2 the corresponding bound is $\lambda'\tau/100\%$. In neither case can the magnitude of this bound be prescribed in planning the amount of observation t or m, since each bound is determined by the outcome of the experiment. The following method gives estimates of prescribed absolute precision.

The problem here is to give an estimate λ' of λ such that with probability at least $1-\alpha$, $|\lambda'-\lambda| \leqq \epsilon$, where α and ϵ are given positive constants. Let n be a positive integer. Observe T_n, the waiting time required for the occurrence of n events. Let $c = \alpha\epsilon^2/2n$. Perform additional observation of the process for $1/2cT_n$ units of time; let X be the number of events observed in this period. Let $\lambda' = 2cT_nX$. Then λ' is an estimate meeting the stated requirements.

The stated properties of λ' may be demonstrated by verifying that λ' has expected value λ and variance $2nc$. Then Tchebycheff's inequality gives

$$\text{Prob}\left\{|\lambda' - \lambda| \leqq \epsilon\right\} \geqq 1 - \frac{2nc}{\epsilon^2} = 1 - \alpha.$$

It remains for further investigation to determine the largest constant c' which can be used to determine the additional amount $1/2c'T_n$ of observation without destroying the desired properties of the estimates $2c'T_nX$. Likewise, rules for the optimal choice of n remain to be investigated.

4. TESTS OF HYPOTHESES ON λ OR θ

A test of a hypothesis specifying the value of λ (and hence of $\theta = 1/\lambda$), say the hypothesis H_0: $\lambda = \lambda_0$, is provided by each method of constructing a confidence interval for λ. If A and B are confidence limits for λ constructed by one of the methods of the preceding sections at confidence level $1-\alpha$, then a test of H_0 at the $1-\alpha$ significance level is obtained by accepting H_0 if $A \leqq \lambda_0 \leqq B$ and otherwise rejecting H_0. Tests of H_0 against one-sided alternatives (e.g., $\lambda > \lambda_0$) can be based on corresponding one-sided confidence limits for λ. Tests with prescribed power can be designed by use of the Poisson or chi-square distributions. As these procedures are relatively standard, they will not be discussed further here.

In each of these testing methods, an economy in the amount of observation required can be achieved by taking advantage of the fact that observation of the Poisson process can be performed continuously. For example, consider tests of H_0: $\lambda = \lambda_0$ against the one-sided alternative $\lambda > \lambda_0$. In the type of test based on a prescribed amount t of observation, the test procedure will amount to rejecting H_0 if the observed number of events x exceeds some "critical number" x_a. Hence if x_a+1 events have been observed at any point in the procedure, observation can be terminated and H_0 rejected at once. In the type of test based on observation of a prescribed number m of events, the test procedure will amount to accepting H_0 if the required amount T of observation is at least equal to some "critical amount" T_a. Hence if m events are not observed before $t = T_a$, observation can then be terminated and H_0 accepted at once, while if m events are observed at or before $t = T_a$, observation can then be terminated and H_0 rejected at once.

Another type of test procedure (which has been investigated by two groups of statisticians, in [8] and [9]) achieves optimal economy in the required amount of observation at the price of a slightly more complicated statistical procedure. Such tests are now available only for testing H_0 against one-sided alternatives. (Two-sided tests of this type can in principle be constructed, by the method used by Wald in [23], pp. 134–37, but this has not yet been done because of certain technical difficulties involved.) These tests can be described as typical sequential probability ratio tests of the kind developed by Wald in [23] and [24], with account and advantage being taken of the fact that the Poisson process may be observed continuously. The typical form of these procedures is the following: Continue observation only so long as the number x of events observed and the amount t of observation performed satisfy the inequality

$$b + st < x < a + st.$$

As soon as $x \leqq b+st$, accept H_0; as soon as $x \geqq a+st$, accept H_1: $\lambda = \lambda_0 + \Delta$. Application of the method may be simplified by a graphical representation of the test inequalities in the (t, x) plane. Here

$$s = \Delta \log\left(\frac{\lambda_0}{\lambda_0 + \Delta}\right), \qquad a = \left(\log\frac{1-\beta}{\alpha}\right)\left(\log\frac{\lambda_0}{\lambda_0 + \Delta}\right),$$

and

$$b = \left(\log\frac{\beta}{1-\alpha}\right)\left(\log\frac{\lambda_0}{\lambda_0 + \Delta}\right),$$

where α is the desired maximum probability of accepting H_1 when $\lambda \leqq \lambda_0$, and β is the desired maximum probability of accepting H_0 when $\lambda \geqq \lambda_0 + \Delta$.

In [10] a number of life testing procedures using Method 2 were investigated, and their operating characteristics and required amounts of test equipment and time determined.

5. COMPARING TWO POISSON PROCESSES

Comparisons of two Poisson processes with parameters λ_1, λ_2 respectively are expressed most usefully in many applications by statements about the value of the ratio $\gamma = \lambda_2/\lambda_1$. The methods described in the following Section 5A provide estimates and tests on γ. In other applications, it may be more appropriate to express comparisons of two processes by statements about the value of the difference $\Delta = \lambda_2 - \lambda_1$; the rather restricted methods available for comparisons in terms of Δ are described in Section 5B. The following examples illustrate possible reasons for choosing either γ or Δ as a criterion of comparison.

Example 1. In order to measure the effectiveness of a shield intended for protection against radiation, we may take any steady source of radiation and a Geiger counter in fixed positions, and record "hits" by radiation alternately with the shield interposed between counter and source and with the shield removed. The natural measure of effectiveness of such a shield is the percentage of radiation it eliminates. If λ_2 is the intensity of radiation at the counter when the shield is present, and λ_1 is the intensity when the shield is removed, then the effectiveness of the shield is $100[1-(\lambda_2/\lambda_1)]\%$. A confidence interval estimate of $\gamma = \lambda_2/\lambda_1$ provides statistical information in an appropriate form, such as, for example, "At confidence level .99, we estimate the effectiveness of the shield to be at least 98.4%."

Example 2. In order to decide which of two samples of ore shows greater radioactivity, as a basis for selecting one of two lots of ore for refining, we can record radiation from each sample with a Geiger counter. If λ_1, λ_2 are the respective intensities of radiation at the counter, the degree to which one lot is preferable to the other is represented by $\Delta=\lambda_2-\lambda_1$, for the gain to be achieved by a correct choice of a better lot is proportional to $|\Delta|$. An appropriate statistical procedure is a test of the hypothesis H_1, that $\lambda_1>\lambda_2$, against the alternative $\lambda_2>\lambda_1$, satisfying the requirement that if $|\Delta|\geqq\Delta'$, where Δ' is the smallest value of $|\Delta|$ of practical significance, then with probability at least $1-\alpha$ the better lot will be chosen.

5A. COMPARISONS IN TERMS OF $\gamma=\lambda_2/\lambda_1$

Method 1. If a prescribed amount t of observation is performed on each of two Poisson processes with parameters λ_1, λ_2, then the numbers x_1, x_2 of events observed in the respective processes will have Poisson distributions with means $\mu_1=\lambda_1 t$, $\mu_2=\lambda_2 t$. Then $\mu_2/\mu_1=\lambda_2/\lambda_1=\gamma$. It has been shown in [19] and [22] that all information on the value of γ contained in the observations x_1, x_2 is obtained by treating the observations as a binomial sample of $m=x_1+x_2$ observations in which x_1 "successes" are observed, and the probability of a "success" is

$$\rho=\frac{\lambda_1}{\lambda_1+\lambda_2}=\frac{1}{1+\lambda_2/\lambda_1}=\frac{1}{1+\gamma}.$$

Thus to test a hypothesis on the value of γ, say H_0: $\gamma=1$ (i.e. $\lambda_1=\lambda_2$), we consider the corresponding test of H_0':

$$\rho=1/2\left(\text{as } \rho=\frac{1}{1+\gamma}=\frac{1}{1+1}\right)$$

when x_1 successes are observed in m binomial trials. We reject H_0 if and only if H_0' is rejected, and at the same significance level. Again, to construct a confidence interval for γ, we can first construct one for ρ (as described in [4]) from our sample of x_1 successes in m trials. If we obtain, for example, "at the .95 confidence level, we estimate that $.33\leqq\rho\leqq.50$," then we can use the equation $\rho=1/(1+\gamma)$, or $\gamma=(1/\rho)-1$, to obtain the equivalent "at the .95 confidence level, we estimate that $1\leqq\gamma\leqq2$."

Method 2. An important weakness of Method 1 is that the procedure cannot be planned to provide any minimum amount of information on γ if the values of λ_1, λ_2 are completely unknown. As we have seen, the

method provides information equivalent to a binomial sample of $m=x_1+x_2$ observations, but m itself is a random observation and may in any particular instance be too small to provide sufficient information (it may even equal zero, and provide no information about γ). Actually m has a Poisson distribution with mean $\mu_1+\mu_2$ so that, if the unknown values of both μ_1 and μ_2 are very small, Method 1 will very often give very little information about γ; even if $\mu_1+\mu_2$ is quite large, there will occur some samples with m too small to give sufficient information about γ.

One way out of this difficulty is apparent. If at least m' binomial observations are required to provide enough information about ρ (and hence about γ) in some application, and if an application of Method 1 gives $m<m'$, another application of Method 1 with an additional amount t_2 of observation of each process may be possible, and if necessary additional amounts of observation until the total number of events observed reaches or exceeds m'. Then the usual binomial statistical methods can be applied to the totals of events observed in each process, as in Method 1.

Method 3. If two Poisson processes can be observed continuously, with observation on each performed at the same rate, then the probability that the ith event observed will occur in the first process is $\rho=1/(1+\gamma)$, for each $i=1, 2, \cdots$. Such observation thus provides a sequence of binomial observations of indefinite length, to which the statistician may apply any of the sequential or curtailed-sampling methods for binomial data (e.g. Wald [23]), or the statistician may terminate observation when a prescribed number n of events has occurred, and apply non-sequential methods for binomial samples of size n. (Methods 2 and 3 were developed in [2].)

Method 4. An alternative way of prescribing observation of two Poisson processes is to require that the first be observed until n_1 events have occurred, and the second until n_2 events have occurred in it. If T_1 is the amount of observation of the first process required to observe n_1 events, then as we have seen above $2\lambda_1 T_1$ has the chi-square distribution with $2n_1$ degrees of freedom; the corresponding quantity for the second process, $2\lambda_2 T_2$, has the chi-square distribution with $2n_2$ degrees of freedom.

The present problem is statistically equivalent to that of comparing variances of two normal populations, given sample variances with $2n_1$ and $2n_2$ degrees of freedom respectively. To test whether $\lambda_1=\lambda_2$, we compute $F=n_2T_1/n_1T_2$, and use the fact that F has the F-distribution with $2n_1$, $2n_2$ degrees of freedom when $\lambda_1=\lambda_2$. To give a confidence in-

terval for $\gamma = \lambda_2/\lambda_1$, we may use the standard method developed for the corresponding problem of giving a confidence interval for the ratio σ_1^2/σ_2^2 of two normal variances by use of the F-distribution. These methods were developed in [11] and [21]. (Cox [5] has investigated the effects of applying to a Method 1 experiment the simpler analysis of Method 4.)

In some applications the present procedure would be more easily applied than those of Methods 2 and 3 above; for example if the processes are separated in space or time the present procedure might be applied. If simultaneous observation of the processes is possible, Methods 2 and 3 would generally give earlier termination since in Method 4 if λ_2 is much smaller than λ_1, T_2 will generally be much larger than T_1.

5B. COMPARISONS IN TERMS OF $\Delta = \lambda_2 - \lambda_1$

Method 1. Let two Poisson processes with parameters λ_1, λ_2 be observed continuously so that the respective waiting times for the first event in each process, T_1 and U_1 say, are obtained, then the waiting times T_2, U_2 for the second event in each process, and so on. We thus obtain a sequence of pairs (T_i, U_i) of waiting times, for each $i = 1, 2, \cdots$; the T_i's have an exponential distribution with mean $\theta_1 = 1/\lambda_1$, the U_i's an exponential distribution with mean $\theta_2 = 1/\lambda_2$. Using such observations we can solve the testing problem stated in Example 2 above.

To do this we apply Girshick's [14] method to the problem of ranking the two exponential population with respect to their means; the resulting test procedure is as follows: After each pair (T_i, U_i) of observed waiting times is obtained, compute

$$Z = \Delta \sum_{i=1}^{m} (T_i - U_i).$$

Continue taking observations until either $Z \geqq a$, in which case accept the hypothesis H_1: $\lambda_2 - \lambda_1 = (1/\theta_2) - (1/\theta_1) \geqq \Delta$, or until $Z \leqq b$, in which case accept the hypothesis H_0: $\lambda_1 - \lambda_2 = (1/\theta_1) - (1/\theta_2) \geqq \Delta$. Here $a = \log (1-\beta)/\alpha$, $b = \log (\beta/1-\alpha)$, where α is the desired maximum probability of rejecting H_0 when H_0 is true, and β the desired maximum probability of rejecting H_1 when H_1 is true. It has been shown in [3] that, unfortunately, Girshick's method can not be generalized to deal with problems like testing whether $\lambda_1 - \lambda_2 \equiv 1/\theta_1 - (1/\theta_2) = 0$ against the alternative $\lambda_2 - \lambda_1 \equiv 1/\theta_2 - (1/\theta_1) = \Delta$.

Method 2. Consider the problem of constructing an estimate Δ^* of $\Delta = \lambda_2 - \lambda_1$, where λ_1, λ_2 are the unknown parameters of two Poisson

processes, such that with probability at least $1-\beta$ $|\Delta^*-\Delta| \leqq \eta$, where β and η are given positive constants. One solution is given by setting $\Delta^*=\lambda_2^*-\lambda_1^*$, where λ_1^* and λ_2^* are obtained as in Method 3 of Section 3 above, taking $\epsilon=\eta/2$ and $(1-\alpha)^2=1-\beta$. In case the two processes can be observed simultaneously, a more efficient solution is the following: Let m be a positive integer. Observe T_m, the waiting time required for the occurrence of a total of m events when the two processes are observed simultaneously. Let $d=\beta\eta^2/2m$. Perform additional observation of each process for $\frac{1}{2}dT_m$ units of time; let Y_1, Y_2 be the respective numbers of events observed in the two processes in this period. Let $\Delta''=2dT_m(Y_2-Y_1)$. Then Δ'' is an estimate meeting the stated requirements. The estimates Δ^* and Δ'' provide useful simple solutions of the common Geiger counter problem of estimating the difference between "noise" count rate and "source plus noise" count rate.

The stated properties of Δ'' are verified as follows: $E(\Delta'')=\Delta$. $E(\Delta''-\Delta)^2=2md$. Then by Tchebycheff's inequality

$$\text{Prob.}\ \{|\Delta''-\Delta| \leqq \eta\} \geqq 1-2md/\eta^2 = 1-\beta.$$

Further investigation should provide rules for optimal choice of n or m, and for shorter periods of additional observation which will suffice to meet the stated requirements.

6. COMPARING THREE OR MORE POISSON PROCESSES

Generalizations of two of the preceding methods for comparing two Poisson processes are available for comparing any number k of processes.

The generalization of Method 4 of Section 5A is the following: Let T_i be the amount of observation required for the observation of a preassigned number n_i of events in the ith process, for $i=1, \cdots k$. Then $2\lambda_i T_i$ has the chi-square distribution with $2n_i$ degrees of freedom, where λ_i is the parameter of the ith process, for each i. Thus this problem is statistically equivalent to that of comparing variances of k normal populations on the basis of sample variances based respectively on $2n_1, \cdots 2n_k$ degrees of freedom. For the problem of testing whether k normal variances are equal, tables of critical values are available in [16]. Let us denote by S_i^2 the sum of squares for the ith normal population, based on $2n_i$ degrees of freedom, used in this test procedure. Then if we substitute for each S_i^2 the observed value T_i, for $i=1, \cdots, k$, and then proceed formally with this test procedure, we obtain a test of the hypothesis that the k Poisson parameters are equal; this test has

the same significance level as the original test, and a corresponding power function.

Every procedure for comparing normal variances can be similarly adapted to give a corresponding procedure for comparing Poisson process parameters.

Method 3 of Section 5A can be generalized as follows: Let k Poisson processes with respective parameters $\lambda_1, \cdots, \lambda_k$ be observed simultaneously. Then the probability that the j^{th} event observed will be observed in the i^{th} process is

$$\rho_i = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_k}, \qquad \text{for each } i.$$

Thus the observed events can be considered independent observations from a multinomial population with class proportions $\rho_1, \cdots, \rho_k$, and all methods for dealing with multinomial data can be applied. For example, the hypothesis that the k processes have equal parameters is equivalent to the hypothesis that the corresponding proportions ρ_i are equal: $H_0: \rho_1 = \cdots = \rho_k = 1/k$. A chi-square test of this hypothesis may be carried out in a $k \times 2$ table; the chi-square statistic for this problem is generally termed the "index of dispersion." The limitations of Method 1 described above apply also to this procedure when a preassigned amount t of observation is used.

Further theoretical discussion of Poisson processes is given in [7] and [12].

REFERENCES

[1] Berkson, Joseph, Magath, Thomas B., and Hurn, Margaret, "Laboratory standards in relation to chance fluctuations of the erythrocyte count as estimated with the hemocytometer," *Journal of the American Statistical Association*, 30 (1935), 414–26.

[2] Birnbaum, Allan, "Some procedures for comparing Poisson processes or populations," *Biometrika*, 45 (1953), 447–49.

[3] Birnbaum, Allan, "Some sequential tests for comparing populations," submitted for publication.

[4] Clopper, C. J., and Pearson, E. S., "The use of confidence or fiducial belts illustrated in the case of the binomial," *Biometrika*, 26 (1934), 403–14.

[5] Cox, D. R., "Some simple approximate tests for Poisson variates," *Biometrika*, 40 (1953), 354–60.

[6] Davis, D. J., "An analysis of some failure data," *Journal of the American Statistical Association*, 47 (1952), 113–50.

[7] Doob, J. L., *Stochastic Processes*, New York, John Wiley and Sons (1953), 98, 398–407.

[8] Dvoretzky, A., Kiefer, J., and Wolfowitz, J., "Sequential decision problems

for processes with continuous time parameter," *Annals of Mathematical Statistics*, 24 (1953), 254–64.

[9] Epstein, B., "Statistical problems in life testing," *Proceedings of the Seventh Annual Convention of the American Society for Quality Control*, May 1953, 385–98.

[10] Epstein, Benjamin, and Sobel, Milton, "Life testing," *Journal of the American Statistical Association*, 48 (1953), 486–502.

[11] Epstein, Benjamin, and Tsao, Chia Kuei, "Some tests based on ordered observations from two exponential populations," *Annals of Mathematical Statistics*, 24 (1953), 458–66.

[12] Feller, William, *An Introduction to Probability Theory and Its Applications*, Vol. I, New York, John Wiley and Sons, Inc. (1950).

[13] Garwood, F., "Fiducial limits for the Poisson distribution," *Biometrika*, 28 (1936), 437–42.

[14] Girshick, M. A., "Contributions to the theory of sequential analysis I," *Annals of Mathematical Statistics*, Vol. 17 (1946), 123–43.

[15] Girshick, M. A., Rubin, H., and Sitgreaves, R., "Estimates of bounded variation in particle counting," Report at September 1952 Meetings of the Institute of Mathematical Statistics.

[16] Hartley, H. O., "Testing the homogeneity of a set of variances," *Biometrika*, 31 (1940), 249–55.

[17] Kitigawa, Tosio, *Tables of the Poisson Distribution*, Tokyo, Japan, Baifukan (1951).

[18] Maguire, B. A., Pearson, E. S., and Wynn, A. H. A., "The time intervals between industrial accidents," *Biometrika*, 39 (1952), 168–80.

[19] Przyborowski, J., and Wilenski, H., "Homogeneity of results in testing samples from Poisson series," *Biometrika*, 31 (1940), 313–23.

[20] Ricker, W. E., "The concept of confidence or fiducial limits applied to the Poisson frequency," *Journal of the American Statistical Association*, 32 (1937), 349–56.

[21] Sukhatme, P. V., "On the analysis of k samples from exponential populations with special reference to the problem of random intervals," *Statistical Research Memoirs*, Vol. I (1936), 94–112.

[22] Tocher, K. D., "Extension of the Neyman-Pearson theory of tests to the discontinuous case," *Biometrika*, 37 (1950), 130–44.

[23] Wald, Abraham, *Sequential Analysis*, New York, John Wiley and Sons, 1947.

[24] Wald, Abraham, "Sequential method of sampling for deciding between two courses of action," *Journal of the American Statistical Association*, 40 (1945), 277–306.

APPLICATIONS OF THE CIRCULAR NORMAL DISTRIBUTION

E. J. GUMBEL
*Columbia University**

IN A previous paper [8] the circular normal distribution was introduced. Now its practical use will be shown. The different procedures explained in the next paragraph will be applied to typical observations taken from geophysical, vital and economic statistics. Finally, a discussion of unsolved problems will be given.

1. PROCEDURE

For simplicity's sake, we consider in the following only circular distributions over the year. The observations are grouped or can be grouped into twelve monthly periods. Then the time series is concentrated in twelve frequencies or relative frequencies p_v ($v = 1, 2 \cdots 12$). In most cases p_v stands for the mean frequency of the vth month, the mean being taken for all observed years. Since the parameters of the circular normal distribution are invariant under a multiplication of all frequencies, it is irrelevant whether the absolute or relative frequencies are used. This important property allows the application of the circular normal distribution to time series which by themselves do not constitute distributions in the usual statistical sense. Let p_v be any non negative numbers of arbitrary nature or dimension, corresponding to the months; then the series $p_v / \sum_1^{12} p_v$ may be regarded as a distribution provided that the summation makes sense and may be analyzed as a circular variate.

However, in our illogical calendar system the lengths of the months vary from 28 to 31 days. If the mode occurs in July, February shows a minimum leading to an apparent asymmetry. Conversely, when the mode occurs in February, its deficiency of two days may create a hole. Since these artificial influences must be eliminated, let the observed frequencies for January, March, May, July, August, October and December be multiplied by $30/31 = 0.96774$ and for February by $30/28 = 1.07143$; let p_v' be the frequencies so adjusted. The year is thus reduced to 360 days. Therefore, the sum of the frequencies is no longer $n = \sum_{v=1}^{12} p_v$.

* Work done in part as Consultant to Stanford University and in part under grant from the Higgins Foundation.

To conserve this observed sum the adjusted frequencies are multiplied by the quotient

$$Q = \frac{p_1 + p_2 + p_3 + p_5 + p_7 + p_8 + p_{10} + p_{12}}{p_1' + p_2' + p_3' + p_5' + p_7' + p_8' + p_{10}' + p_{12}'}.$$

In this procedure the observed frequencies for April, June, September, and November are preserved while the frequencies for the other months are twice adjusted. The adjustments must be used if the differences between the frequencies are small, say of the order 10%, since they are then strongly affected by the different lengths of the months. The adjustments are not relevant if the differences between different months are large, say if the largest monthly frequency is 10 times the smallest one. The second adjustment is unnecessary if the quotient Q differs from unity by less than 1%.

The adjusted frequencies are attributed to the 15th of each month and may be traced as a linear histogram or on polar paper consisting of concentric equidistant circles and radii for each degree. The maximum frequency is traced at north and the following months are traced clockwise. After the choice of a unit distance the square roots of the twice adjusted frequencies, $\sqrt{p_\nu''}$, are plotted instead of the adjusted frequencies p_ν'' themselves. This procedure equalizes the areas of the observed and the theoretical distributions. Distributions differing with respect to the sample size n are thus traced in different scales. However, a *uniform scale* is obtained if all adjusted frequencies $\bar{p}''$ are divided by their mean $\bar{p}$. This method has the advantage that distributions with different values of n may be traced on the same scale. It has the disadvantage that the frequencies are no longer visible.

In the use of circular normal distribution

$$\phi(\alpha) = \frac{e^{k \cos(\alpha - \alpha_0)}}{2\pi I_0(k)}, \tag{1.1}$$

the mode α_0 is estimated from

$$\tan \alpha_0 = \sum_1^{12} \sin \alpha_\nu \Big/ \sum_1^{12} \cos \alpha_\nu \tag{1.2}$$

where

$$\sum_1^{12} \cos \alpha_\nu = \text{July} - \text{Jan.} + 0.86603(\text{Aug.} - \text{Feb.} + \text{June} - \text{Dec.}) + 0.5(\text{Sept.} - \text{Mar.} + \text{May} - \text{Nov.}), \tag{1.3}$$

$$\sum_{1}^{12} \sin \alpha_\nu = \text{Oct.} - \text{Apr.} + 0.86603(\text{Sept.} - \text{Mar.} + \text{Nov.} - \text{May}) + 0.5(\text{Aug.} - \text{Feb.} + \text{Dec.} - \text{June}). \tag{1.4}$$

and the months are written instead of their frequencies p_ν''. This determines α_0 only up to 180°. The exact location is found from the conventional diagram for the signs of the trigonometric functions. In the estimation of α_0 it is sufficient to calculate to whole degrees of the angle.

The parameter k is estimated from

$$\bar{a} = \frac{1}{n}\sqrt{\left(\sum_{1}^{12} \sin \alpha_\nu\right)^2 + \left(\sum_{1}^{12} \cos \alpha_\nu\right)^2} \tag{1.5}$$

with the help of Table II[1] which gives k as function $\bar{a}$. The parameters for the reduced values $p_\nu''/\breve{p}$ are obtained in an analogous manner. The adjustments for months of equal lengths may be introduced into (1.3) and (1.4). This leads to

$$\begin{aligned}\sum_{1}^{12} \cos \alpha_\nu = {} & 0.96677(\text{Jul.} - \text{Jan.}) + 0.86603(\text{June}) \\ & + 0.83726(\text{Aug.} - \text{Dec.}) + 0.5(\text{Sept.} - \text{Nov.}) \\ & + 0.48339(\text{May} - \text{Mar.}) - 0.92789\ \text{Feb.},\end{aligned} \tag{1.6}$$

$$\begin{aligned}\sum_{1}^{12} \sin \alpha_\nu = {} & 0.96677\ \text{Oct.} + 0.86603(\text{Nov.} + \text{Sept.}) \\ & + 0.48339(\text{Aug.} + \text{Dec.}) - \text{Apr.} \\ & - 0.83726(\text{Mar.} + \text{May}) - 0.53571\ \text{Feb.} - 0.5\ \text{June}.\end{aligned} \tag{1.7}$$

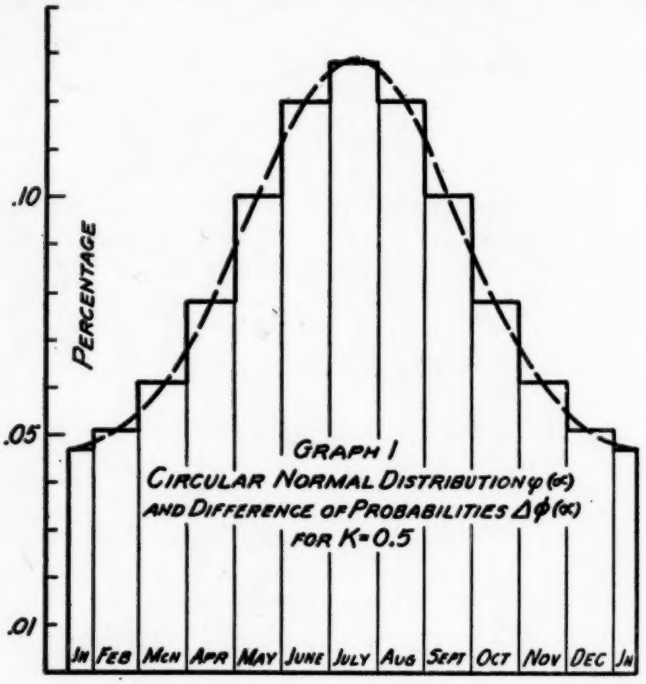
Since the parameters α_0 and k are invariant under a multiplication of the frequencies no adjustment is then necessary to conserve the sum of the observations. However, for the comparison to the theory the observations must again be adjusted for months of equal lengths and for the conservation of the observed sum. Therefore, this system is not used in the examples that follow.

In another analytic procedure a constant value equal to the minimum monthly value $\breve{p}$ is assumed to hold throughout the year. This component is treated as a circular uniform distribution. The other component consisting of the observed frequencies $p_\nu - \breve{p}$ is considered as a circular normal distribution with $n' = n - 12\breve{p}$. However, up to

[1] Tables with Roman numbers refer to the previous publication [8].

now no case has occurred in which this procedure gives a better fit than the previous one.

The theoretical values corresponding to the observations $\sqrt{p_v}$ are obtained after multiplying the values given in Table III by 0.28868 $\sqrt{n}$. If k is nearly equal to one of the values given in Table II, it is sufficient to use a slide rule. The theoretical values corresponding to reduced frequencies $p_v/\bar{p}$ are taken from III without any multiplication.



GRAPH I
CIRCULAR NORMAL DISTRIBUTION $\varphi(\alpha)$
AND DIFFERENCE OF PROBABILITIES $\Delta\phi(\alpha)$
FOR K=0.5

The theoretical values are plotted on the polar diagram as points. A continuous symmetrical curve is obtained by joining the points to the left and right of the mode by the same parts of a French ruler. If the modal direction is practically zero or a multiple of 30 degrees, i.e. if the observed mode is either conserved or shifted by these amounts, we may use differences of probabilities $\Delta\Phi(\alpha)$ instead of the densities $\phi(\alpha)$. This procedure leads to a comparison of the observed with the theoretical wedge diagram. It has the advantage that the conventional criteria for the goodness of fit can be used.

The differences of probabilities (Table 1) were obtained by taking

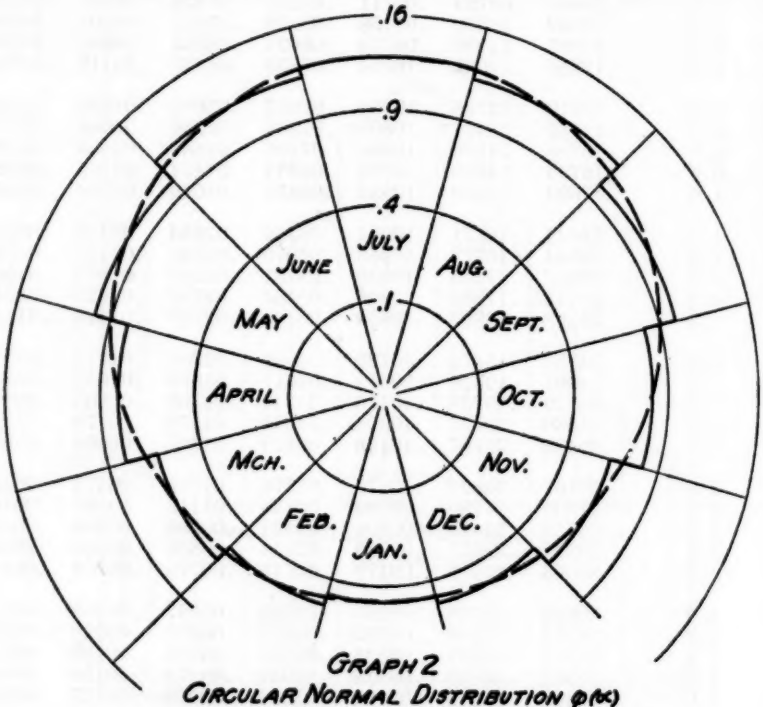
TABLE 1

THE CIRCULAR NORMAL PROBABILITY DIFFERENCES $\Delta\Phi(\alpha)$ FOR 12 AREAS, OF 30°, CENTERED ON:

k	Mean	±30°	±60°	±90°	±120°	±150°	±180°
0.0	.08333	.08333	.08333	.08333	.08333	.08333	.08333
0.1	.09176	.09056	.08735	.08314	.07912	.07630	.07530
0.2	.10054	.09793	.09111	.08255	.07476	.06953	.06770
0.3	.10962	.10539	.09458	.08158	.07031	.06304	.06058
0.4	.11895	.11286	.09774	.08024	.06581	.05690	.05394
0.5	.12846	.12032	.10054	.07858	.06133	.05110	.04780
0.6	.13810	.12768	.10296	.07662	.05690	.04570	.04217
0.7	.14782	.13491	.10501	.07440	.05256	.04069	.03704
0.8	.15755	.14196	.10666	.07196	.04837	.03608	.03239
0.9	.16726	.14880	.10793	.06933	.04434	.03186	.02822
1.0	.17690	.15539	.10882	.06656	.04049	.02804	.02449
1.1	.18644	.16171	.10934	.06369	.03686	.02459	.02118
1.2	.19584	.16772	.10953	.06076	.03345	.02149	.01826
1.3	.20507	.17345	.10938	.05781	.03026	.01872	.01569
1.4	.21413	.17885	.10895	.05484	.02730	.01627	.01344
1.5	.22298	.18395	.10824	.05190	.02457	.01410	.01149
1.6	.23163	.18874	.10729	.04901	.02206	.01219	.00979
1.7	.24007	.19320	.10613	.04617	.01978	.01052	.00833
1.8	.24829	.19738	.10476	.04344	.01768	.00906	.00707
1.9	.25630	.20127	.10324	.04078	.01578	.00778	.00599
2.0	.26409	.20487	.10158	.03822	.01407	.00668	.00506
2.1	.27167	.20822	.09979	.03578	.01252	.00572	.00427
2.2	.27905	.21130	.09790	.03345	.01113	.00489	.00360
2.3	.28623	.21415	.09592	.03124	.00988	.00418	.00303
2.4	.29322	.21677	.09388	.02914	.00876	.00356	.00255
2.5	.30003	.21917	.09179	.02715	.00776	.00304	.00214
2.6	.30666	.22138	.08966	.02528	.00687	.00258	.00179
2.7	.31312	.22338	.08752	.02352	.00607	.00220	.00150
2.8	.31942	.22522	.08535	.02186	.00536	.00187	.00126
2.9	.32557	.22688	.08318	.02031	.00474	.00158	.00105
3.0	.33157	.22838	.08100	.01886	.00418	.00135	.00088
3.1	.33744	.22974	.07884	.01750	.00369	.00114	.00073
3.2	.34317	.23096	.07670	.01624	.00325	.00096	.00061
3.3	.34878	.23204	.07458	.01505	.00286	.00082	.00051
3.4	.35427	.23301	.07249	.01394	.00252	.00069	.00043
3.5	.35964	.23386	.07042	.01292	.00222	.00058	.00036
3.6	.36490	.23460	.06839	.01196	.00196	.00049	.00030
3.7	.37006	.23524	.06639	.01107	.00172	.00042	.00025
3.8	.37513	.23579	.06442	.01026	.00151	.00035	.00020
3.9	.38009	.23625	.06251	.00949	.00133	.00030	.00017
4.0	.38497	.23662	.06063	.00877	.00117	.00025	.00014
4.1	.38976	.23692	.05879	.00812	.00102	.00021	.00012
4.2	.39446	.23714	.05700	.00750	.00090	.00018	.00010

half of the successive differences of the probability function calculated to 9 decimals and attributing them to the midpoints of the intervals.

Graphs 1 and 2 show a wedge diagram obtained from Table 1 and the continuous distribution traced on linear and on aequiareal polar scales. Since the observations can only be traced in the wedge form, while the theoretical distribution can also be traced as a continuous curve, it is



GRAPH 2
CIRCULAR NORMAL DISTRIBUTION $\phi(x)$
AND DIFFERENCE OF PROBABILITIES $\Delta\phi(x)$
TRACED ON AN EQUIVALENT POLAR SCALE
$K=0.5$

desirable to compare the appearance of one and the same distribution in the two forms. The graphs also show the type of deviations that are inevitable if an observed wedge diagram is compared to a continuous distribution. With the exception of the modal and anti-modal months, the curves representing the theoretical distribution intersect the observed wedge diagram near the middle of each month and the wedge diagram falls short (exceeds) the theoretical curve at the beginning (end) of each month. The tracing of $\Delta\Phi(\alpha)$ is a graphical alternative to the tracing of the distribution $\phi(\alpha)$. In general it has the disad-

vantage that the "months" considered therein do not coincide with the months in the calendar.

The different analytical and graphical procedures outlined will now be applied to numerical data.

2. GEOPHYSICAL OBSERVATIONS

The following examples deal with rainfall, run-off and evaporation, i.e. the hydrologic cycle and temperatures. The *amounts of rainfall per month* form a genuine distribution. For the Esopus watershed which is essential to the water supply of New York City, Table 2 shows the mean monthly rainfall. The modal (anti-modal) month is written in the first (last) line. The second column contains the observed numbers in inches, the third introduces the adjustments for length of months. The sum of the adjusted frequencies is 30.665″ while the observed sum for the same months before adjustment is 31.36″. Consequently the adjusted frequencies are multiplied by $31.36/30.665 = 1.02263$, resulting in the fourth column which conserves the observed total mean annual rainfall $n = 48.91''$. The square roots of the numbers in the fourth column starting with July are plotted in the wedge diagram, Graph 3. The variations indicate a systematic circular behavior starting with a mode in July and decreasing to the anti-mode in January with the exception of the rainfalls in November which exceed those in October.

TABLE 2

MEAN MONTHLY RAINFALL IN INCHES: ESOPUS WATERSHED

1 Month	2 Obs.	3 First Adj.	4 Sec. Adj.	1 Month	2 Obs.	3 First Adj.	4 Sec. Adj.
July	4.69	4.539	4.64	—	—	—	—
June	4.54	—	4.54	Aug.	4.58	4.432	4.53
May	4.31	4.171	4.27	Sept.	4.50	—	4.50
Apr.	4.23	—	4.23	Oct.	4.03	3.900	3.99
Mar.	3.85	3.726	3.81	Nov.	4.28	—	4.28
Feb.	3.05	3.268	3.34	Dec.	3.45	3.339	3.42
—	—	—	—	Jan.	3.40	3.290	3.36

To obtain the theoretical distribution we first calculate the parameters. Formulas (1.3) and (1.4) lead from Table 2, column 4 to the sums

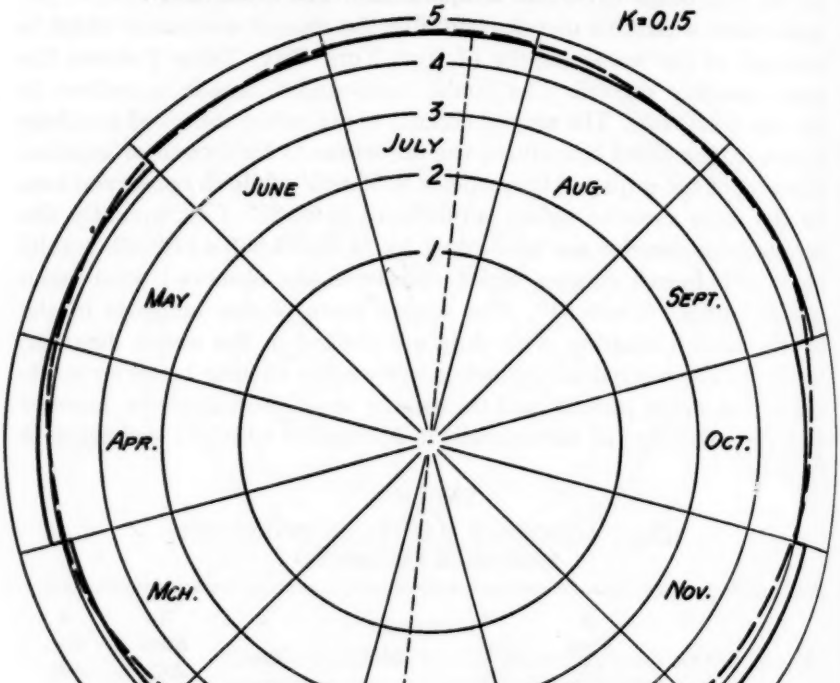
$$\sum_1^{12} \cos \alpha_\nu = 3.62553; \qquad \sum_1^{12} \sin \alpha_\nu = 0.40112.$$

Consequently

$$\tan \alpha_0 = 0.11067; \qquad \alpha_0 = 6^0.$$

Since α is counted from 15 July, this means a shift of the mode to 21 July. It is due to the fact that the frequencies for the latter part of the

GRAPH 3
MONTHLY MEAN RAINFALL ESOPUS WATERSHED



year show a slight predominance over the first part (20.72″ against 20.19″). Formula (1.5) leads to

$$\bar{a} = 3.64766/48.91 = 0.07458.$$

Table II gives the values

$$\bar{a} = 0.07, \quad k = 0.1403; \qquad \bar{a} = 0.08, \quad k = 0.1605.$$

Linear interpolation leads to $k=0.1495$. It is sufficiently accurate to choose $k=0.15$. This small value of the parameter is due to the fact that the differences between the frequencies are quite small.

From Table III we obtain radii vectors $\psi(\alpha)\sqrt{12/n}$ by interpolating between $k=0.1$ and $k=0.2$. These values are given in Table 3, column 2. Multiplication by $\sqrt{48.91/12}=2.0189$ leads to the radii vectores (column 3), which are plotted on the original grid of the polar paper to the left and to the right of the modal value at 21 July. The fit of the theory to the observations leaves nothing to be desired.

TABLE 3

CALCULATION OF RADII VECTORS

1	2	3
Angle	$\psi(\alpha)\sqrt{12/n}$	$\psi(\alpha)$
0°	1.075	2.170
20°	1.070	2.160
40°	1.051	2.122
60°	1.035	2.090
80°	1.012	2.043
100°	0.9840	1.987
120°	0.9603	1.939
140°	0.9414	1.901
160°	0.9293	1.876
180°	0.9252	1.868

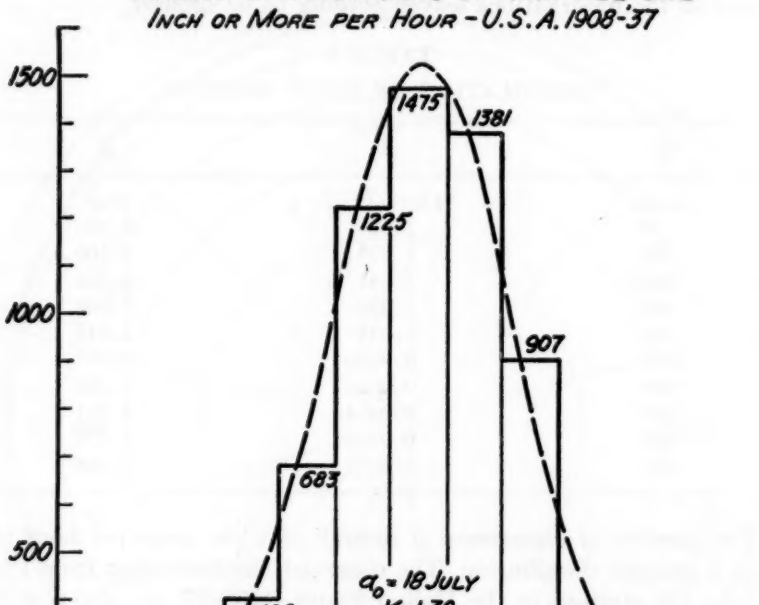
The *numbers of occurrences of rainfall* of 1″ or more per hour also form a genuine distribution. The observed numbers taken from Dyck [7] for 156 stations in the United States, 1908–37, are given at the bottom of the linear Graph 4. The sum of the adjusted frequencies is 4366.5 while the corresponding sum without adjustment is 4502. Multiplication of the adjusted numbers by 1.03103 conserves the observed total sum 7235. The twice adjusted numbers are given at the top of Graph 4. The adjustments are less than 1% except for February where the observed value is decreased by more than 10%. The mode in July is more than 14 times the minimum in January. Therefore, the parameter k is relatively high, $k=1.34$. The calculation of the parameters in this and all following examples is given in Table 9. The distribution is slightly skewed to the right; consequently, the mode is shifted from 15 to 18 July.

The theoretical distribution is obtained by the same procedure as before. We interpolate $\psi(\alpha)\sqrt{12/n}$ from Table III, multiply these volumes by $\sqrt{n/12}=24.554$ to obtain $\psi(\alpha)$. The squares of these num-

bers are plotted in Graph 4 and traced in a continuous curve.

The *monthly runoff* in inches for the watershed of the Derwent River at Yorkshire Bridge, Derbyshire (England) taken from [14], p. 397, is given in Table 4 for the 43 years June 1905 to May 1948. The monthly percentages of the annual runoff form a distribution. Therefore the monthly runoff can be analyzed as a cyclical variate.

GRAPH 4
NUMBER OF OCCURANCES OF RAINFALL ONE INCH OR MORE PER HOUR - U.S.A. 1908-'37



The second column gives the runoff for the 43 years. The third column gives the monthly means adjusted for length of months, the fourth conserves the observed yearly sum $n=36.28''$. The distribution is slightly asymmetrical since the mode is in January while the antimode is in June. The observations and the theory are traced in equivalent polar and in linear scales in Graph 5. The fit is quite satisfactory, since the deviations between theory and observations do not show any

systematic trend. The authors of the article [14] assumed that the differences between the months are due to chance. This opinion implies a uniform distribution of the monthly runoff which seems highly improbable, to say the least.

The circular normal distribution cannot be applied to rivers with two regimes, one derived from rainfalls, say in Autumn, and the other from the melting of snow, say in Spring. These conditions lead to a bi-modal distribution.

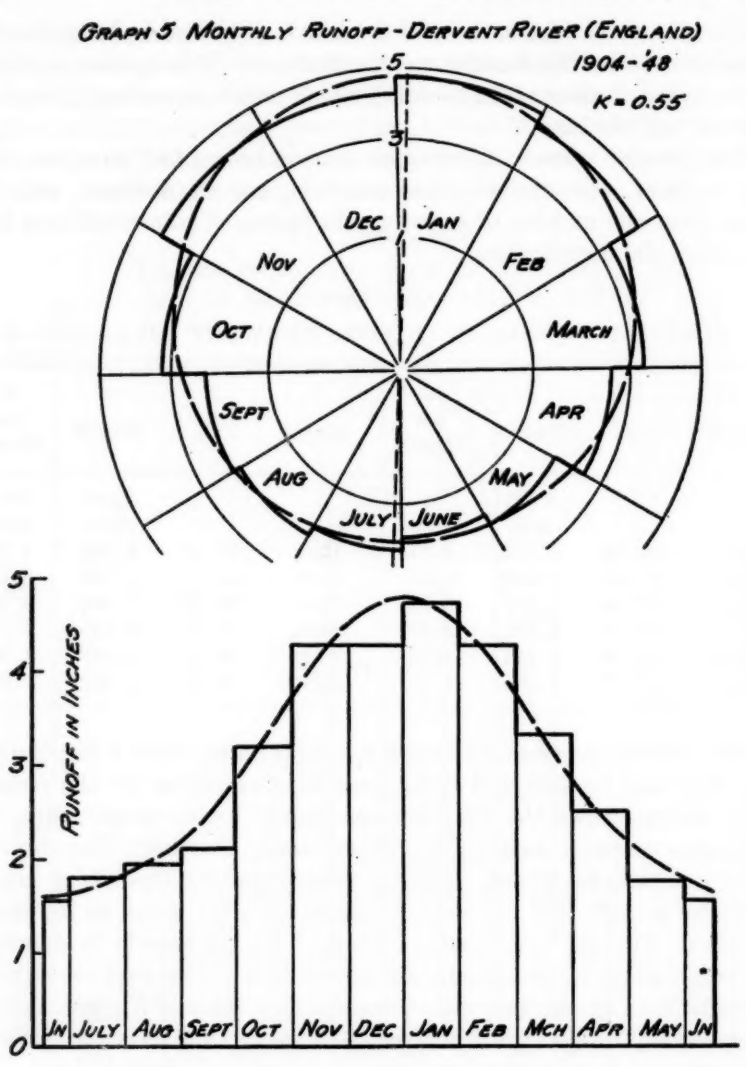
TABLE 4

MONTHLY RUNOFF IN INCHES: DERWENT RIVER 1905–48

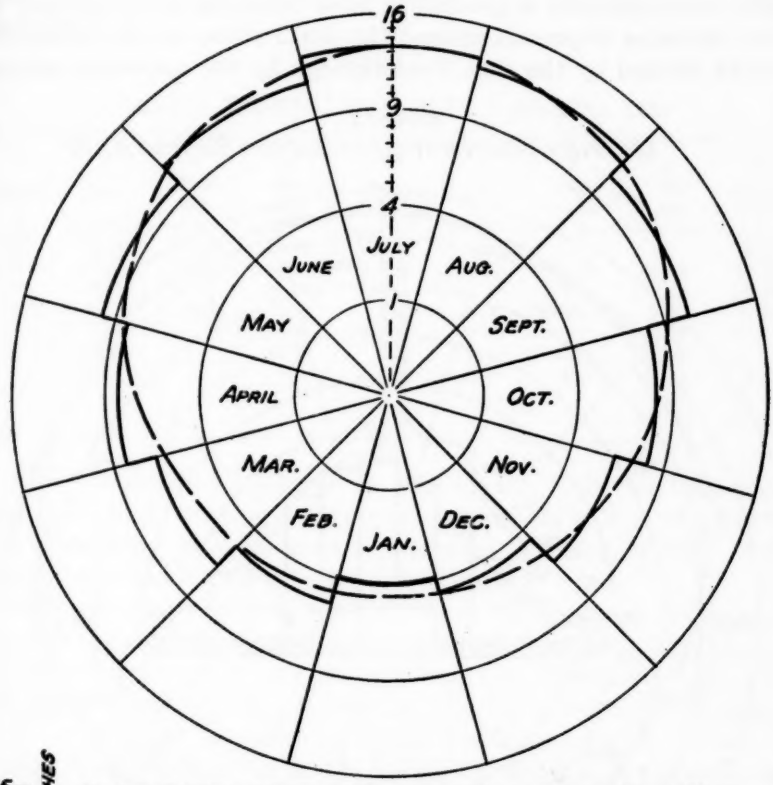
1 Month	2 Sum	3 Runoff	4 per Month	1 Month	2 Sum	3 Runoff	4 per Month
		First Adj.	Sec. Adj.			First Adj.	Sec. Adj.
Jan.	210.29	4.733	4.81	Dec.	191.47	4.309	4.38
Feb.	171.87	4.282	4.35	Nov.	185.17	4.306	4.31
Mar.	147.85	3.327	3.38	Oct.	142.28	3.202	3.26
Apr.	106.84	2.485	2.49	Sept.	91.57	2.130	2.13
May	78.85	1.775	1.80	Aug.	86.54	1.948	1.98
June	67.27	1.564	1.56	July	80.09	1.802	1.83

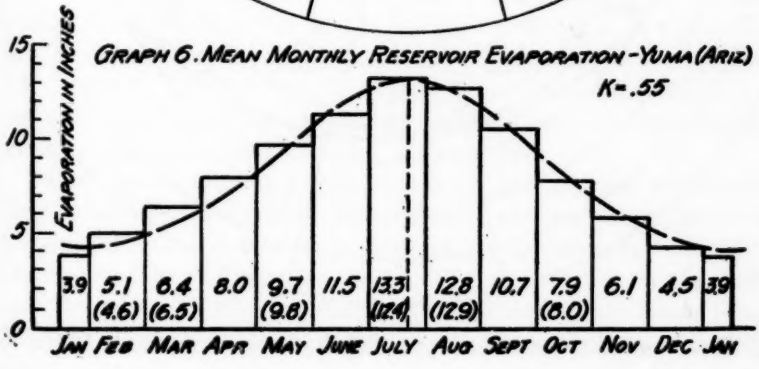
The *monthly evaporations* from a reservoir do not form a distribution but they may be analyzed as the preceding examples for the reasons given above. Since the evaporations depend upon temperature, the maximum occurs in summer and the minimum in winter. The observations in inches for Yuma, Arizona, taken from the Hydrology Handbook [9] p. 127, and the twice adjusted data are given in the linear Graph 6. The mode occurs in July and the anti-mode in January. However, the evaporations in the second half of the year are slightly stronger than in the first one. Consequently there is a slight shift in the mode to 21 July. The theory leads to a very good fit to the observations.

The *mean monthly temperatures* at a given place do not constitute a distribution, since temperatures cannot be added. However, it is certainly legitimate to trace them as a function of time. This is done in Graph 7 for the mean monthly temperatures in Boston (Mass.) taken from Conrad [2] p. 110. A reduction for months of equal lengths does not seem to be necessary. The circular distribution, used to reproduce the data, gives an excellent picture.



We conclude from the previous examples, which are illustrative of a wide range of problems, that the circular normal distribution is an efficient tool for the analysis of meteorological phenomena, even in cases which are not actually distributions.

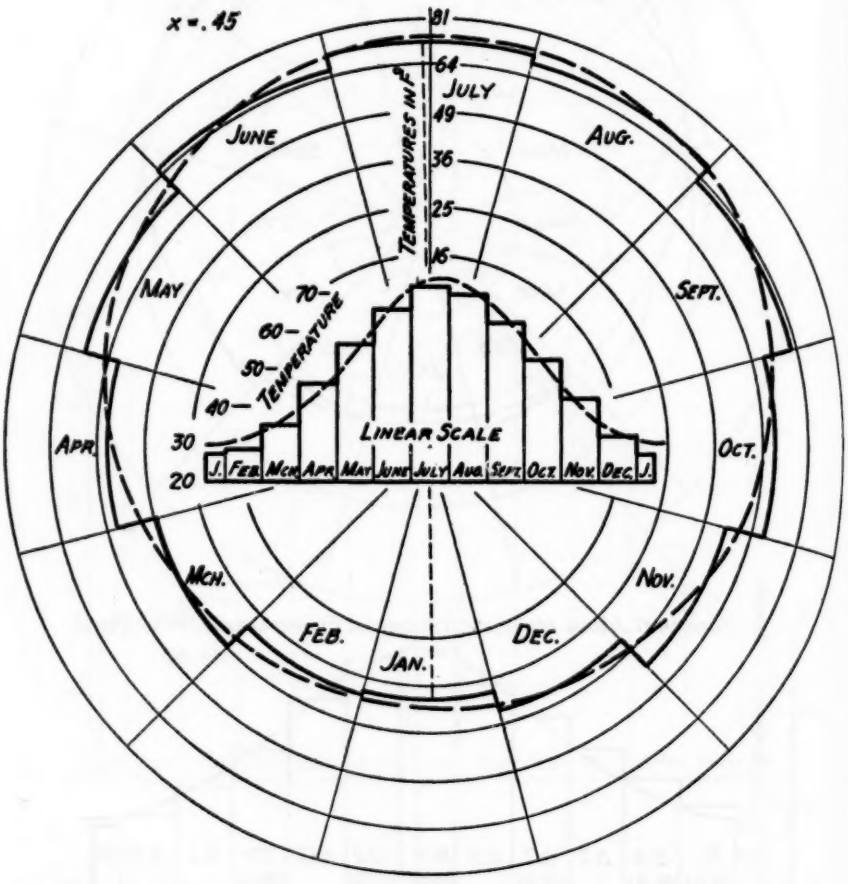

16
9
4
JULY
AUG.
SEPT.
OCT.
NOV.
DEC.
JAN.
FEB.
MAR.
APRIL
MAY
JUNE


GRAPH 6. MEAN MONTHLY RESERVOIR EVAPORATION - YUMA (ARIZ)
K=.55
EVAPORATION IN INCHES
15
10
5
0
3.9
5.1
(4.6)
6.4
(6.5)
8.0
9.7
(9.8)
11.5
13.3
(12.4)
12.8
(12.9)
10.7
7.9
(8.0)
6.1
4.5
3.9
JAN FEB MAR APR MAY JUNE JULY AUG SEPT OCT NOV DEC JAN

3. APPLICATIONS TO VITAL STATISTICS

The percentage of persons who die each month forms a distribution. Since the population is growing in most countries it is customary in vital statistics to present instead the death rates, i.e. the number of deaths divided by the time lived through by the respective popula-

GRAPH 7
MONTHLY TEMPERATURES - BOSTON (Conrad p.110)



tion. These *monthly death rates* per 1000 have the dimension of the reciprocal of the time and can be compared month by month, and year by year. They may be treated for our purposes as if they constituted a distribution, since the death rates are proportional to the number of deaths for a stationary population.

The monthly means of the observed and adjusted death rates in the United States, September 1946 to August 1951 are given in Table 5, where the mode and anti-mode are given in the first and last lines respectively.

TABLE 5

DEATH RATES U. S. A. SEPT. 1946–AUG. 1951

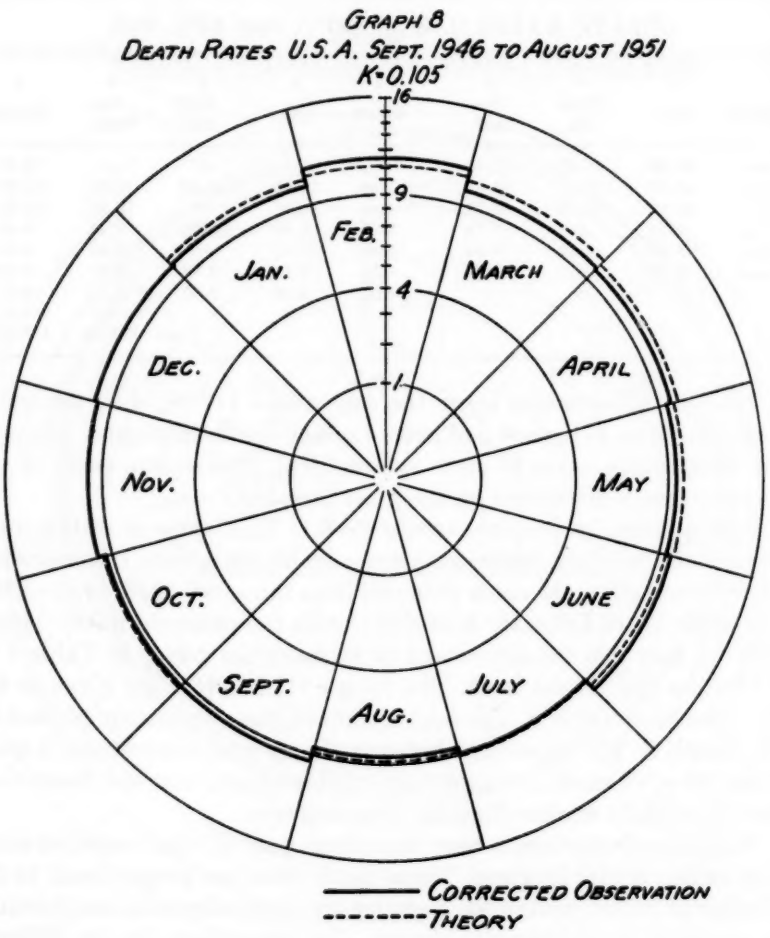
1 Month	2 Obs.	3 First Adj.	4 Sec. Adj.	1 Month	2 Obs.	3 First Adj.	4 Sec. Adj.	5 Theory
Feb.	10.66	11.42	11.64	—	—	—	—	10.82
Jan.	10.50	10.16	10.35	Mar.	10.80	10.45	10.65	10.68
Dec.	10.42	10.08	10.27	Apr.	10.24	—	10.24	10.31
Nov.	9.70	—	9.70	May	9.60	9.29	9.46	9.81
Oct.	9.42	9.12	9.29	June	9.42	—	9.42	9.34
Sept.	9.08	—	9.08	July	9.18	8.88	9.05	9.00
—	—	—	—	Aug.	8.96	8.67	8.83	8.88
							$n=$ 117.98	117.98

The two adjustments leave the sum at $n=117.98$, shift the mode from March to February and leave the anti-mode in August. Thereby the distribution is made more symmetrical. The square roots of the corrected rates are traced on the polar Graph 8.

It is sufficiently accurate to use $k=0.1$. This value is rather small because the monthly variations are also small. Since both trigonometric sums are negative, the mode obtained from $\tan \alpha_0 = 0.58487$ is $\alpha_0 = 210°$. The mode on 15 February is identical with the observed mode. Therefore, we may use the differences of probabilities $\Delta\Phi(\alpha)$ in Table 1 to obtain the theoretical rates. The values $117.98\Delta\Phi(\alpha)$ are given in the last column of Table 5. The square roots of these values are plotted on the Graph 8. The agreement between theory and observation is quite close. No systematic deviations are visible except that the theoretical mode is slightly smaller than the observed one.

Infant death rates are a very important part of vital statistics since they reflect Social Hygiene. These death rates are proportional to the number of infant deaths for a stationary population and can therefore be analyzed by the circular theory. The mean rates for the different months twice adjusted are shown at the bottom of the linear Graph 9. The mode and the decrease of the frequencies toward the anti-mode are quite systematic: the monthly variations are very small. Consequently, the value of the parameter k also turns out to be very small and the theoretical distributions traced in Graph 9 resemble a uniform distribution.

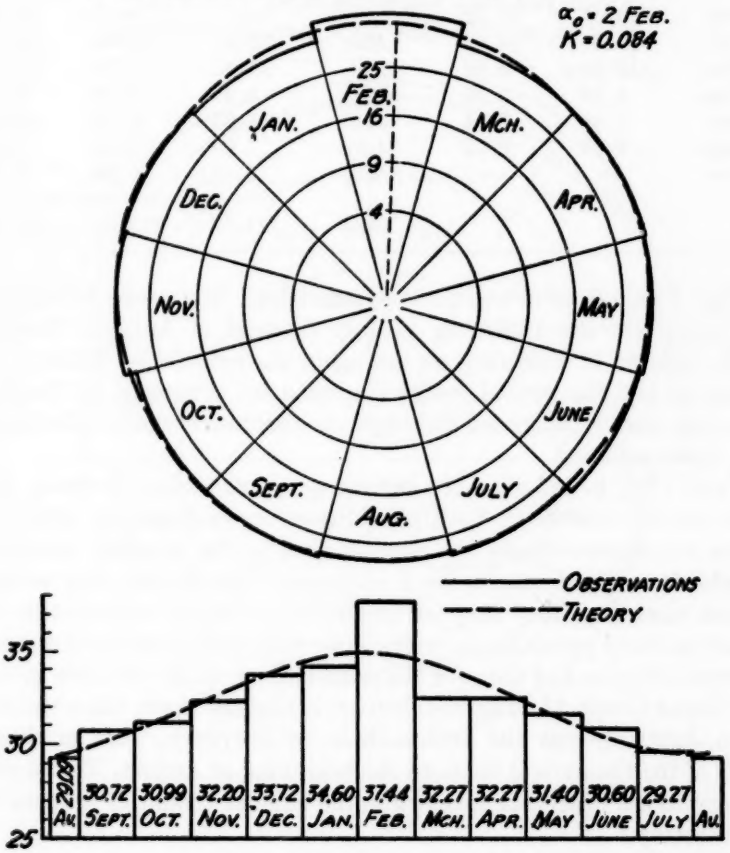
The cyclical variation of the number of deaths over the years is a very complex result of different causes of deaths having maxima at different times. For example, influenza has a mode in January, pneumonia in February, scarlet fever and meningitis in March, measles in



April, dysentery in July, typhoid fever in August, poliomyelitis in September, diphtheria in October, and tularemia in December. The day with maximum probability of deaths should be considered as critical. The popular idea of the existence of critical days for certain illnesses is completely legitimate.

The mean percentages of *deaths of children* under 1 year of age in the United States, August 1947 to July 1951, due to *pneumonia and influenza* are given in Table 6. These percentages are proportional to the number of deaths from these causes for a constant number of

GRAPH 9
INFANT DEATH RATES U.S.A. SEPT. 1945 TO AUG. 1951



children who died. Therefore they may be represented by the circular distribution. The sum of the observations is $n = 114.27$; the sum of the corrected data is $n' = 113.32$. Since the second adjustment would have amounted to less than 1%, it was not applied.

TABLE 6

PERCENTAGE OF DEATH OF CHILDREN DUE TO PNEUMONIA AND INFLUENZA, U. S. A. 1947–1951

1 Month	2 Frequency Observed	3 Adjusted	1 Month	2 Frequency Observed	3 Adjusted	4 Theor. Rate
Feb.	16.07	17.22	—	—	—	15.38
Jan.	14.49	14.02	Mar.	13.97	13.52	14.27
Dec.	10.31	9.98	Apr.	11.46	11.46	11.60
Nov.	9.40	9.40	May	8.35	8.08	8.74
Oct.	7.58	7.34	June	6.05	6.05	6.57
Sept.	6.02	6.02	July	5.05	4.89	5.33
—	—	—	Aug.	5.52	5.34	4.93
			Sum	114.27	113.32	113.33

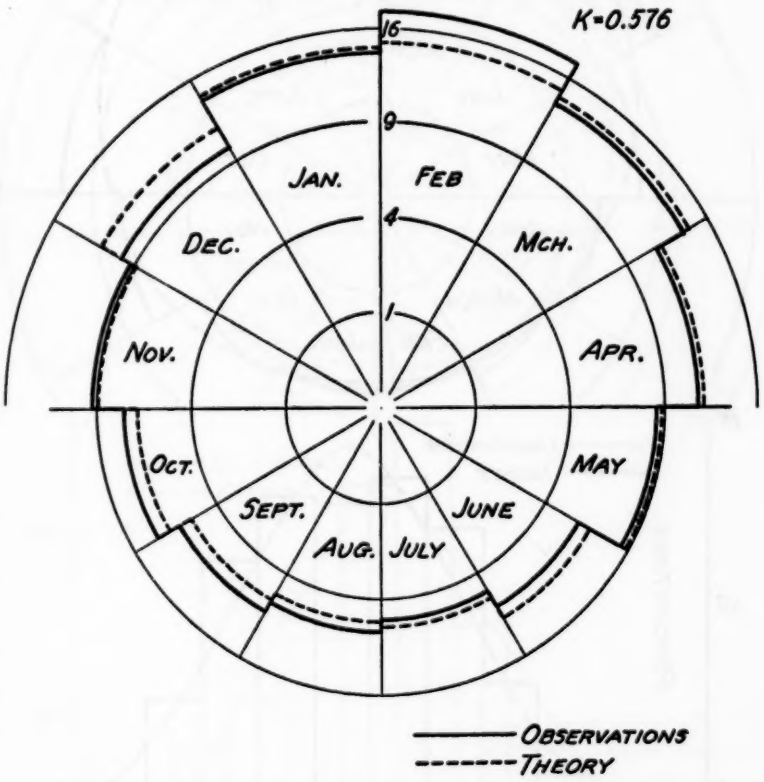
The distribution is not quite symmetrical. The mode is located in February but the anti-mode in July (instead of August). Since the mode falls on 15 February we use again the probability Table 1. The observed and theoretical wedge diagrams are compared in Graph 10. The first may be accepted although the theoretical mode falls short of the observed one.

Pearl [12], p. 153 gives the *percentage of total infant mortality due to diarrhea and enteritis* in the United States for each month, 1933–1935. Since these percentages are proportional to the monthly number of deaths from these causes for a stationary population with constant infant mortality, they may be treated as a circular distribution. The mean monthly percentages for the three years adjusted first for months of equal lengths and then for the conservation of the sum are given in the linear Graph 11. The distribution is slightly skew, since the mode is in July, whereas the anti-mode is in February. This produces a shift of the theoretical mode to the beginning of August. The fit of the theory to the observation shown in the circular Graph 11 is again very satisfactory.

The observed and twice adjusted *numbers of persons drowned in each month* in the United States in 1946 are given respectively in the bottom and in the end of the linear Graph 12. The square roots of the corrected numbers are traced in the wedge diagram. Evidently the maximum must lie somewhere in the summer. However, some unfortunate persons drown even in winter by breaking through the ice.

The circular character of the distribution is thus clear. However, the distribution is not quite symmetrical. The mode is in July, but the anti-mode is in December. The frequencies for the beginning of the year are stronger than for the end. Therefore, the theoretical mode is at 26 June instead of 4 July where it should be according to popular customs. The mode of the theoretical curve is smaller than the observed

GRAPH 10
PERCENT OF PNEUMONIA AND INFLUENZA
IN INFANT DEATHS U.S.A. AUG. 1947 TO JULY 1951

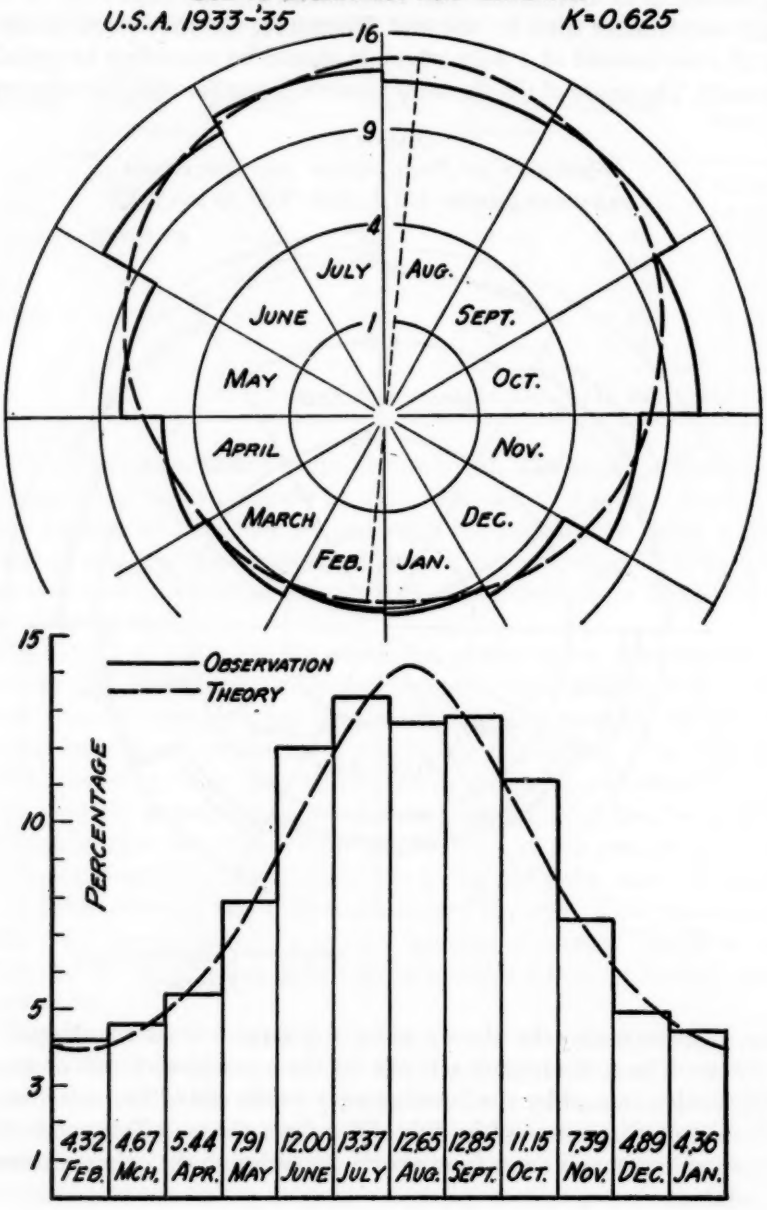


one. Consequently the theory gives too large values for August to October. These deviations are due to the skewness of the observed distribution caused by the heterogeneity in the data. The activities in which drowning occurred in July differ from those in December. The drowned in summer come from another population than those drowned in winter.

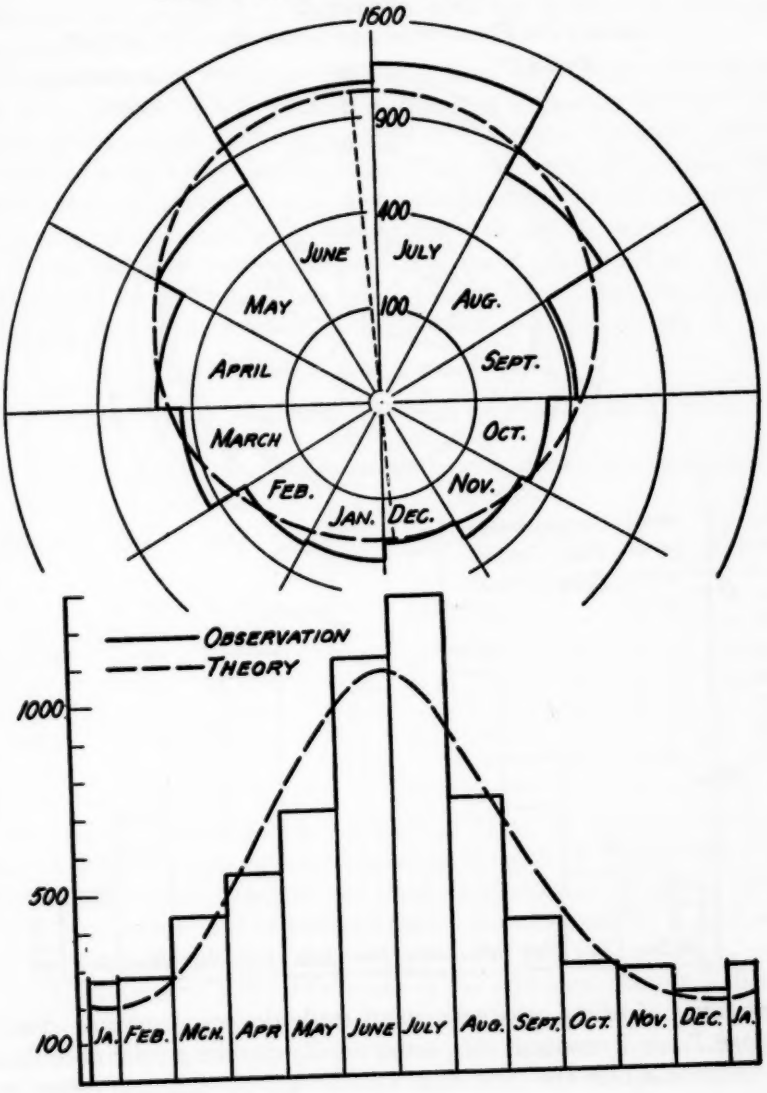
GRAPH II

PERCENT OF TOTAL INFANT MORTALITY DUE TO DIORRHEA AND ENTERITIS

U.S.A. 1933-'35 K=0.625

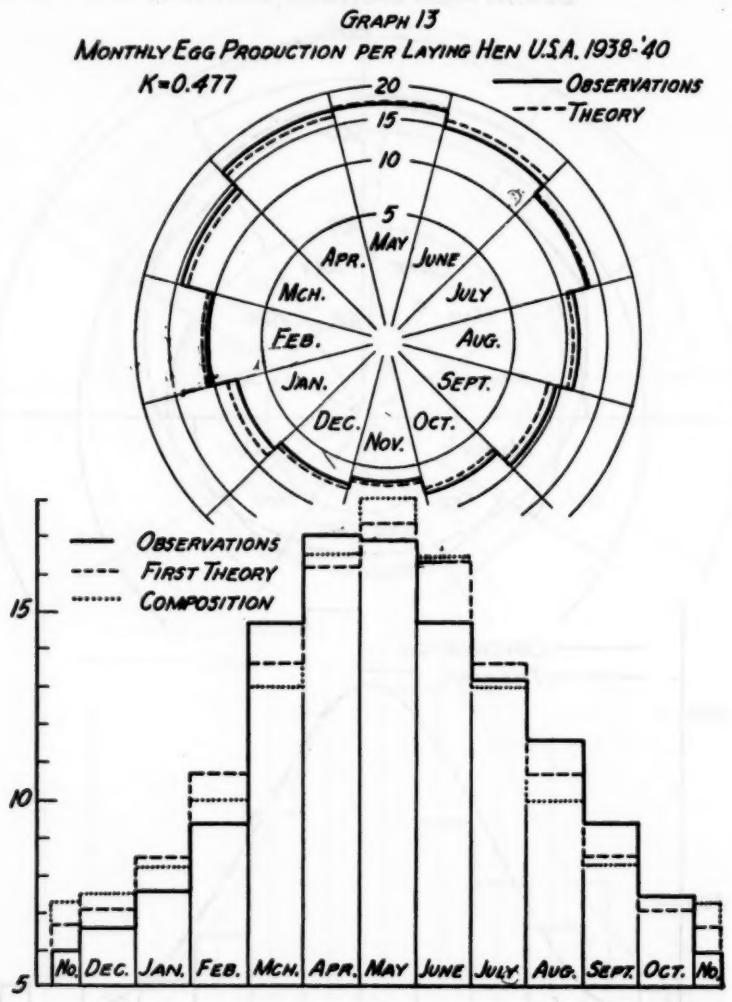


GRAPH 12
DEATHS FROM DROWNING, U.S.A. 1946



4. APPLICATIONS TO ECONOMIC TIME SERIES

Other possible applications may be found in economic time series. But they entail certain difficulties which must be pointed out. First,



most series of prices, production and trade do not constitute distributions. Second, economic time series usually contain secular or cyclical movements, which the economist would want to eliminate before investigating the seasonal movements; but the process of eliminating these undesirable influences and expressing the series as a set of sea-

sonal relatives, as it is often done, leaves the series with less resemblance to a distribution than it had originally. Finally, the seasonals in many time series have configurations that cannot be adequately fitted by the assumption of a single circular distribution. These three difficulties are amplified in connection with the illustrative examples presented below.

The monthly *egg production per laying hen*, Table 7, given by Kendall [10], p. 368, is not, strictly speaking, a frequency distribution. Although the eggs produced in a period of years by a single hen, a flock of hens or even by all hens consist of discrete countable units, the process of dividing the number of eggs produced by the number of hens producing them conceals the actual egg frequencies, which are not available. However, the eggs-to-hens ratio may be regarded as a sort of distribution for an average hen or an average flock. The observed and twice adjusted data are given in Table 7 and traced in Graph 13.

TABLE 7

MONTHLY EGG PRODUCTION PER LAYING HEN U. S. A. 1938–1940

1 Month	2 Obs.	3 Adj.	1 Month	2 Obs.	3 Adj.	4 First Theory	5 Second Theory
May	17.10	16.91	—	—	—	17.20	18.13
Apr.	17.00	17.00	June	14.77	14.77	16.14	16.47
March	14.90	14.73	July	13.40	13.25	13.59	12.99
Feb.	9.53	10.43	Aug.	11.77	11.63	10.74	10.01
Jan.	7.70	7.61	Sept.	9.47	9.47	8.48	8.29
Dec.	6.67	6.66	Oct.	7.60	7.51	7.13	7.52
—	—	—	Nov.	6.03	6.03	6.69	7.30

The calculations lead to a mode on 16 May. We choose instead 15 May, a difference which is not visible in the graphs. Interpolation for $k=0.477$ in Table 1 leads to the theoretical number of eggs given in Table 7, column 4. The fit between theory and observations is good and no systematic deviations exist.

If we consider only the differences of the numbers given in Table 7, column 3 from their minimum 6.033, the sum of these numbers is $n'=63.64$. The resulting vector strength obtained from (1.5) is $\bar{a}=0.4961$. The direction of the mode remains unchanged. Table II leads to the estimate $k=1.14$. The theoretical results $6.03+63.64\Delta\Phi(\alpha)$ obtained from Table 1 are given in the last column of Table 7 and are

traced in the lower part of Graph 13. The difference in the two representations are not visible in the polar Graph 13. The distribution obtained by the composition of a uniform and a circular distribution does not show a better fit to the observations than the simple procedure used before.

The observations in the following examples are treated according to the uniform procedure outlined in paragraph 1.

The series on *automobile production* [1] is a genuine frequency distribution consisting of discrete, countable units—i.e., the number of automobiles produced in three years allocated by months. Since these are original data, no adjustments have been made for cyclical, secular or irregular influences, except those for length of months (column 3).

TABLE 8

PASSENGER CARS PRODUCTION

1 Month	2 Obs.	3 Adj.	4 Redcd.	1 Month	2 Obs.	4 Adj.	4 Redcd.
Aug.	1589	1538	1.193	—	—	—	—
Sept.	1453	1453	1.127	July	1435	1389	1.078
Oct.	1523	1674	1.144	June	1527	1527	1.185
Nov.	1251	1251	0.971	May	1197	1158	0.898
Dec.	1192	1154	0.895	Apr.	1200	1200	0.931
Jan.	1120	1084	0.841	Mar.	1223	1184	0.919
—	—	—	—	Feb.	985	1055	0.818
				Sum	15695	15667	12.000

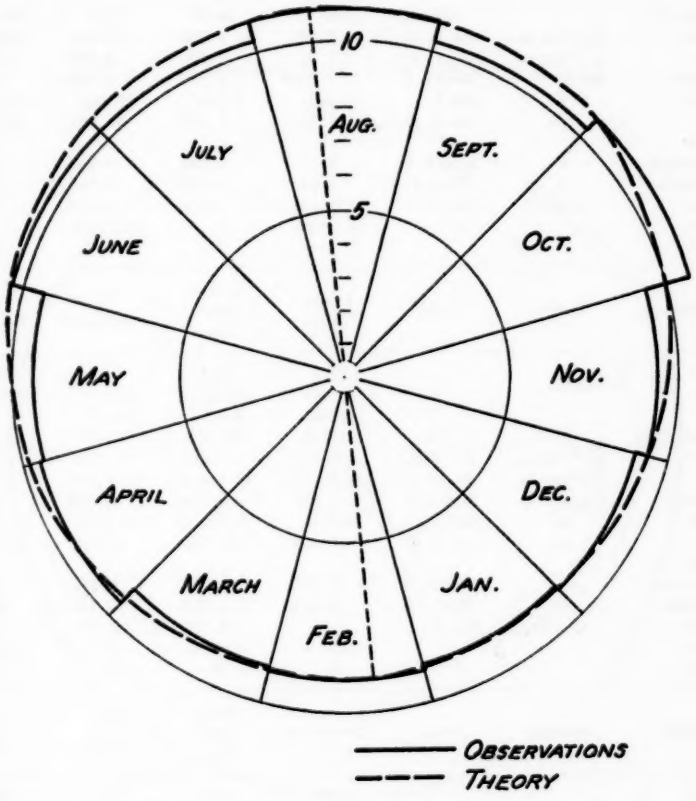
The mean monthly number of cars produced is 1308. Division of column 3 by this number leads to column 4, the sum of which is of course 12.

The irregular movements are rather pronounced. August is the modal month, but October and June exceed September and July, hence these months appear as conspicuous valleys. Since $n=12$ the values $\psi(\alpha)$ in Table III constitute the theoretical distribution. It is sufficiently accurate to use $k=0.2$. The smooth curve in Graph 14 does not give a bad fit and especially the mode and anti-mode are conserved.

The series on *pig-iron* production, taken from Croxton and Cowden [5], p. 374, consists of seasonal relatives corrected for cyclical, secular and irregular influences. The resulting wedge diagram, Graph 15 is, therefore, fairly regular. The amplitude is not very large. The anti-mode precedes the mode by 5 instead of by 6 months, implying a some-

what skewed distribution. The mode is well reproduced and on the whole the fit is good, since the deviations between theory and observations do not show any systematic pattern.

GRAPH 14
PASSENGER CAR PRODUCTION U.S.A. 1948-'50 IN UNIFORM SCALE



In the *sales of Misses' coats*, 1926–30 [11] p. 393, two modes and two antimodes exist, April, October and July, February, respectively. Clearly such observations cannot be fitted into the scheme of a single circular distribution.

5. OUTLOOK

The introduction of a new distribution solves certain statistical problems, but at the same time raises a considerable number of new problems, some of which are enumerated below:

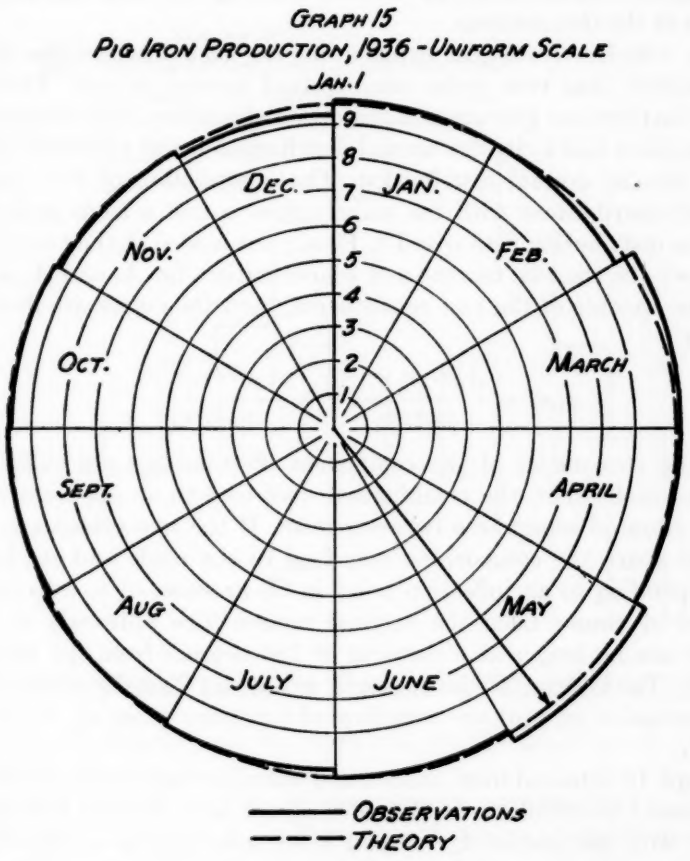
TABLE 9
ESTIMATION OF PARAMETERS

1 Material	Symbol	From	Rainfall	Runoff	Evaporation	Temperature
2 Place			USA 156 Stations	Derwent	Yuma (Ar.)	Boston (Mass.)
3 Date			1908–37	1905–48		
4 Sample Size	n	Obs.	7235	36.28	99.9	595.2
5 Cosine	$\sum_1^{12} \cos \alpha_s$	eq. (1.3)	3997.23	− 9.355	26.081	130.19
6 Sine	$\sum_1^{12} \sin \alpha_s$	eq. (1.4)	237.31	2.086	2.856	21.58
7 Tangent	$\tan \alpha_s$	lines 5, 6	0.0594	− 0.223	0.1095	0.1657
8 Mode	α_s	line 7	18 July	2 Jan.	21 July	24 July
9 Sum	$n\bar{s}$	eq. (1.5)	4004.27	9.548	26.237	131.96
10 Vector Strength	$\bar{s}$	lines 4, 9	0.553	0.264	0.263	0.222
11 Parameter	k	Table II	1.338	0.548	0.545	0.455
12 Factor	$\sqrt{n/12}$	line 4	24.554	1.739	2.885	7.043
13 Table			in graph	4	in graph	—
14 Graph			4	5	6	7

1 Material	Symbol	From	Death Rate	Infant Death Rate	% Pneumonia	% Diarrhea
2 Place			USA	USA	USA	USA
3 Date			Se 46–Au 51	Se 45–Au. 51	Au. 47–Ju. 51	1933–1935
4 Sample Size	n	Obs.	117.98	383.97	113.32	101.00
5 Cosine	$\sum_1^{12} \cos \alpha_s$	Eq. (1.3)	− 5.385	−15.241	−27.232	26.732
6 Sine	$\sum_1^{12} \sin \alpha_s$	eq. (1.4)	− 3.082	− 4.545	−15.507	12.934
7 Tangent	$\tan \alpha_s$	lines 5, 6	0.572	0.298	0.569	0.484
8 Mode	α_s	line 7	15 Feb.	2 Feb.	15 Feb.	11 Aug.
9 Sum	$n\bar{s}$	eq. (1.5)	6.204	15.746	31.338	29.696
10 Vector Strength	$\bar{s}$	lines 4, 9	0.053	0.041	0.276	0.294
11 Parameter	k	Table II	0.105	0.082	0.576	0.625
12 Factor	$\sqrt{n/12}$	line 4	—	5.657	3.071	2.901
13 Table			5	in graph	6	in graph
14 Graph			8	9	10	11

1 Material	Symbol	From	Drowning	Eggs per Hen	Car Prod.	Pig Iron
2 Place			USA	USA	USA	USA
3 Date			1946	1938–40	1948–50	1936
4 Sample Size	n	Obs.	6632	136	12.00	12.00
5 Cosine	$\sum_1^{12} \cos \alpha_s$	eq. (1.3)	2398	16.502	1.045	0.3635
6 Sine	$\sum_1^{12} \sin \alpha_s$	eq. (1.4)	−850	−26.915	0.494	− 0.4756
7 Tangent	$\tan \alpha_s$	lines 5, 6	−0.354	−1.631	0.473	−1.308
8 Mode	α_s	line 7	26 June	16 May	10 Aug.	22 May
9 Sum	$n\bar{s}$	eq. (1.5)	2450	31.572	1.156	0.599
10 Vector Strength	$\bar{s}$	lines 4, 9	0.384	0.232	0.096	0.050
11 Parameter	k	Table II	0.832	0.477	0.194	0.100
12 Factor	$\sqrt{n/12}$	line 4	23.509	3.367	1	1
13 Table			in graph	7	8	—
15 Graph			12	13	14	15

If the deviations between the observations and a circular uniform distribution are small, a criterion for the reality of a cycle is needed, see [6]. We want to know whether a mode and anti-mode exist or are spurious and due to chance. In statistical terms this means: let the population value of the parameter be κ (e.g. zero); what is the probability of obtaining a certain value of k or a larger one from a sample of



size n? A similar, but more complicated, problem is to test the differences between the modes and values of the parameter k obtained from two samples. Finally a test for the goodness of fit between theory and observations is needed.

The following systematic deviations between theory and observations may exist: the observed values about the mode exceed (fall short

of) the theoretical values and show the opposite behavior at the antimode; the observed mode differs sensibly from the theoretical one; the observed antimode is not distant 180° from the observed mode. Finally there may be two modes and two antimodes. One of the latter may not be visible. In all these cases the chance distribution about one fixed date does not explain the observations. Then there must exist systematic reasons which have to be investigated from the intrinsic nature of the observations.

The simplest statistical explanation for such divergencies is the assumption that two cycles are involved instead of one. Two uniform distributions give again a uniform distribution. The composition of a uniform and a circular normal distribution leads, as shown before, to a circular normal distribution. The composition of two circular normal distributions with the same values α and k leads again to a circular distribution with α and k. Finally let $\alpha_1 \neq \alpha_2$ be the two modes, let $k_1 \neq k_2$ be the two measures of concentration, let A_1 and A_2 be the relative weights of the two constituents, then the composite distribution is

$$\phi(\alpha) = \frac{A_1 e^{k_1 \cos(\alpha - \alpha_1)}}{2\pi I_0(k_1)} + \frac{A_2 e^{k_2 \cos(\alpha - \alpha_2)}}{2\pi I_0(k_2)} .$$

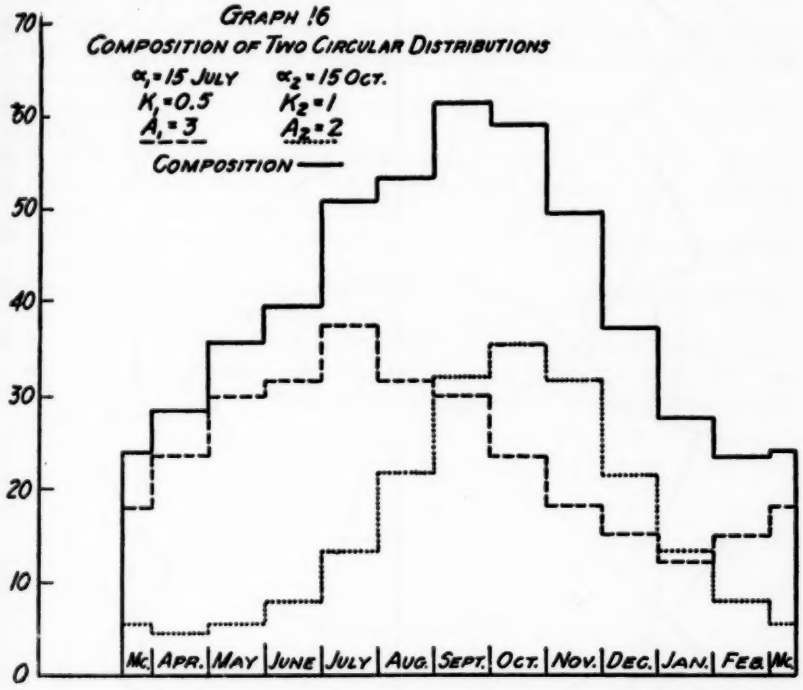
If the two modes of the component distributions are sufficiently near to each other, the combination may lead to an apparent single mode situated somewhere between them. If the two modes are sufficiently apart, the composition may lead to one mode and one hump, corresponding to an inflection point in linear scale, or to two modes, shifted of course from the original modes. The antimode or antimodes are no longer at a distance of 180 degrees from the mode or modes. The location of the modes or antimodes formally obtained by differentiation of $\phi(\alpha)$ are complicated functions of α_1, α_2, k_1, k_2 and A_1, A_2.

Graph 16 obtained from Table 1 and traced to linear scale, shows two component distributions with modes α_1 and α_2 at 15 July and 15 October, with parameters $k_1 = .5$ and $k_2 = 1$ and weights $A_1 = 3$, $A_2 = 2$. They combine into an asymmetrical distribution with a mode in September and a minimum in February.

Graph 17 traced in aequiareal scale shows the composition of two circular distributions of equal weight and with values $k = 2$, centered on $\alpha_1 = 0$ $\alpha_2 = \pi$ together with the two components.

The converse problem of separating an observed asymmetrical or symmetrical circular distribution into two symmetrical parts presents

no logical difficulties. However, great analytic difficulties arise from the estimation of the five parameters α_1, α_2, k_1, k_2, where $A_1=A$, $A_2=1-A$. Karl Pearson [13] has solved the corresponding problem within the linear normal distribution. In the general case it leads to an equation of the ninth degree. The difficulties are reduced see [3], if certain assumptions can be made about the modes or the values of the parameters k or the weights. By this procedure asymmetrical



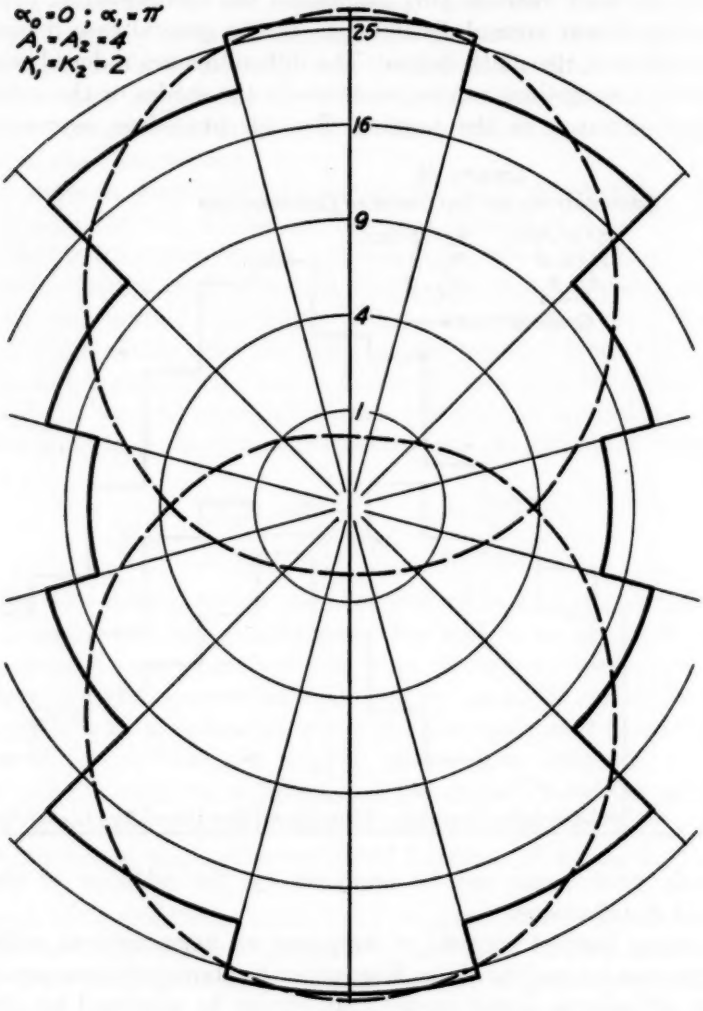
periodic phenomena can be analyzed by the addition of circular normal distributions.

Another method consists of wrapping an asymmetrical unlimited distribution around the circle. Even after the above problems are solved there will remain many cycles that cannot be analyzed by circular distributions. Statistics were not statistics if everything could be explained by this method.

Acknowledgment: The author is indebted to Arnold Court (University of California, Berkeley), David Durand (National Bureau of Economic Research), and J. A. Greenwood (Manhattan Life Insurance Co.) for valuable suggestions.

GRAPH 17

COMPOSITION OF TWO CIRCULAR DISTRIBUTIONS



BIBLIOGRAPHY

[1] *Automobile Facts and Figures;* Automobile Manufacturers Association, 31st edition, New York (1951), 5.

[2] Conrad, V. A., *Methods in Climatology*, Cambridge, Mass. (1946), chapter VIII.

[3] Court, A., "Separating frequency distributions into two normal components," *Science*, 110 (1949), 24.

[4] Court, A., "Some new statistical techniques in geophysics," *Advances in Geophysics*, edited by H. L. Landsberg, vol. 1, New York (1952), 75–83.
[5] Croxton, F. F., and Cowden, D. J., *Applied General Statistics*, New York (1940), 374.
[6] Durand, David, and Greenwood, J. A., "The integral solution of Pearson's random walk problem and related matters," *Annals of Mathematical Statistics*, 24 (1953), 686–87.
[7] Dyck, H. D., and Mattice, W. A., "A study of excessive rainfalls," *Monthly Weather Review*, 69 (1941), 293–302.
[8] Gumbel, E. J., Greenwood, J. A., and Durand, David A., "The circular normal distribution: theory and tables," *Journal of the American Statistical Association*, 48 (1953), 131–52.
[9] *Hydrology Handbook*, American Society of Civil Engineers, New York, (1949), 127.
[10] Kendall, M. G., *The Advanced Theory of Statistics*, London (1946), 127.
[11] Kuznetz, Simon, *Seasonal Variations in Industry and Trade*, National Bureau of Economic Research, New York (1933), 368.
[12] Pearl, R., *Introduction to Medical Biometry and Statistics*, Philadelphia (1940), 153.
[13] Pearson, K., "On the dissection of frequency curves," *Philosophical Transactions*, Royal Society, 185 A (1894), 71–110.
[14] Thompson, R. W. S., "The application of statistical methods in the determination of the yield of a catchment from run-off data," with a statistical note by D. H. Thomson, *Journal of the Institution of Water Engineers*, 4 (1950), 397.

A NEW TYPE OF CONTROL CHART LIMITS FOR MEANS, RANGES, AND SEQUENTIAL RUNS

H. WEILER

New South Wales University of Technology, Sydney, Australia

Since most conventional control chart limits are designed so that the ratio of the expected number of false alarms to the number of samples tested is fixed in advance irrespective of the sample size, the ratio of false alarms to the number of *articles* tested varies with the sample size. In this paper, control charts are designed for which the expected ratio of false alarms to the number of articles tested is independent of the sample size. The power of the control charts in relation to sample size is investigated.

1. INTRODUCTION AND SUMMARY

THE usual control chart controlling the mean of a normal population is constructed in the following way: After the mean and standard deviation of the population have been reliably estimated, samples of fixed size n are selected and their arithmetic means $\bar{x} = \Sigma\, x/n$ are calculated. A chart is then constructed with control limits $m \pm B\sigma/\sqrt{n}$, where m and σ are estimates of the population mean and standard deviation, and B a constant. The various values of $\bar{x}$ are entered in the chart in chronological order, and as soon as one such value falls outside the control limits, production is stopped to allow investigation.

It is customary to take B equal to 3 or 3.09, irrespective of the sample size n. If $B = 3.09$ and if the population mean and standard deviation remain unchanged, an average of 500 samples will be required to produce one $\bar{x}$ value above the upper or below the lower limit. This means that an average of $500n$ articles will be tested for every one false alarm raised. The usual practice of setting control limits entails therefore that the average number of articles tested between two false alarms depends on the sample size n used for the control chart. This is likely to be an undesirable feature in quality control, where the cost of inspection is usually proportional to the number of articles inspected. The production engineer will therefore be interested in the number of articles rather than the number of samples tested.

In this paper, tables are provided for the determination of control limits for which the average number of false alarms (or Type I errors) is a fixed percentage of the number of articles tested, *independent of the sample size*. Curves similar to power curves are drawn for various sample sizes, giving the average amount of inspection in terms of the

amount by which the population mean has changed. A similar procedure is adopted for range charts.

One important result of the investigations is that the power of a mean chart increases rapidly with the sample size. For instance, the average amount of inspection required to detect a given change of the population mean is in most practical cases about twice as large for sample size $n=5$ as it would be if a chart for sample size $n=10$ were used (unless the change of the mean is very large). For samples of 20, the average amount of inspection can, under certain favorable circumstances be as little as one quarter of the amount required with samples of 5. This has already been pointed out in two previous papers [2, 3]; the results obtained here are even more striking.

However, since it is often necessary to maintain small samples in spite of the loss of power, it was shown in [3] that the power of small samples can be improved by charts using sequential runs. In the present paper, tables similar to those mentioned above are provided for the determination of suitable control limits for run charts. The corresponding curves show that the amount of inspection is greatly reduced, although it is frequently still high compared with the amount required for simple large sample charts.

Finally, run charts for ranges are briefly investigated. It turns out that the power of range charts is practically independent of the sample size and that the use of runs does not represent an improvement.

2. CONTROL LIMITS FOR MEAN CHARTS

If P is the probability that a random sample mean falls above the upper control limit, then about $100P$ samples in every 100, or one in every $1/P$ samples will fall above the upper control limit. It follows that on the average n/P *articles* will be tested before an alarm is raised at the upper limit.

Now, let p be the probability that a random sample causes a *false* alarm at the upper limit, and let a be the average number of articles tested before the false alarm is raised. We have then $a=n/p$ or $p=n/a$. If n and a are given, p can be determined; and if we assume x to be normally distributed, we can determine B such that the probability is p that $\bar{x}$ exceeds the upper control limit $m+B\sigma/\sqrt{n}$, where m is the mean and σ the standard deviation of the parent population. If, for instance, we take $n=5$ and $a=5000$, we have $p=0.001$, and we obtain $B=3.09$ from a set of normal tables.

Proceeding in this manner for various sample sizes, but keeping $a=5000$ fixed, we obtain the values of p, B, $B/\sqrt{n}$ shown in Table I.

TABLE I

VALUES OF p, B, $B/\sqrt{n}$, WHEN $a=5000$

Notation: p = probability that a random sample causes a false alarm; n = sample size; a = average number of articles tested before an alarm occurs at the upper (or lower) control limit, if the process remains under control.

n	3	4	5	6	8	10
p	.0006	.0008	.0010	.0012	.0016	.0020
B	3.24	3.16	3.09	3.03	2.95	2.88
$B/\sqrt{n}$	1.87	1.58	1.38	1.24	1.04	0.91

Similar arguments hold for the lower control limit, so that in each case an average of one false alarm above the upper and one false alarm below the lower control limit must be expected for every 5000 articles tested. In other words, we will be able to test an average of 2500 articles before one false alarm is raised either way.

Control limits for values of a other than 5000 may be calculated in the same way. Table II gives values of $B/\sqrt{n}$ for various values of a and for sample sizes ranging from 3 to 50.

TABLE II

VALUES OF $B/\sqrt{n}$ FOR CONTROL LIMITS $m \pm (B/\sqrt{n})\sigma$ FOR MEAN CHARTS

Notation: n = sample size; m = population mean; σ = population standard deviation; a = average number of articles tested before an alarm occurs at the upper (or lower) control limit, if the process remains under control.

n	$a=1000$	$a=2000$	$a=3000$	$a=4000$	$a=5000$
3	1.59	1.71	1.78	1.83	1.87
4	1.33	1.44	1.50	1.54	1.58
5	1.15	1.26	1.31	1.35	1.38
6	1.02	1.12	1.17	1.21	1.24
7	0.93	1.02	1.07	1.10	1.13
8	0.85	0.94	0.99	1.02	1.04
9	0.79	0.87	0.92	0.95	0.97
10	0.74	0.82	0.86	0.89	0.91
15	0.560	0.628	0.665	0.690	0.710
20	0.459	0.520	0.553	0.577	0.593
25	0.392	0.448	0.479	0.500	0.515
30	0.343	0.396	0.425	0.444	0.458
35	0.306	0.356	0.383	0.402	0.416
40	0.277	0.324	0.350	0.368	0.381
45	0.253	0.299	0.324	0.340	0.353
50	0.232	0.277	0.301	0.317	0.329

To determine the control limits $m \pm (B/\sqrt{n})\sigma$, we estimate the mean m and standard deviation σ of the population and obtain the appropriate factor $B/\sqrt{n}$ from Table II. If, for instance, we want to use samples of size $n=8$, and if we wish to test an average of 2000 articles before raising a false alarm either way ($a=4000$), we find $B/\sqrt{n}=1.02$, and the control limits are $m \pm 1.02\sigma$.

3. THE AVERAGE AMOUNT OF INSPECTION WHEN THE POPULATION MEAN CHANGES

If the population mean changes from m to $m+k\sigma$ ($k>0$) while σ remains constant, the variate $\bar{x}$ will have the mean $m+k\sigma$ and the standard deviation $\sigma/\sqrt{n}$, so that the variate

$$z = \frac{\bar{x} - (m + k\sigma)}{\sigma/\sqrt{n}} \tag{1}$$

is a standardized normal variate (mean zero and S.D. one). The probability that $\bar{x}$ exceeds the upper control limit $m+B\sigma/\sqrt{n}$ is then

$$\begin{aligned} P &= \Pr\left\{\bar{x} \geqq m + \frac{B\sigma}{\sqrt{n}}\right\} = \Pr\left\{\bar{x} - m - k\sigma \geqq \frac{B\sigma}{\sqrt{n}} - k\sigma\right\} \\ &= \Pr\left\{\frac{\bar{x} - m - k\sigma}{\sigma/\sqrt{n}} \geqq B - k\sqrt{n}\right\} = \Pr\{z \geqq B - k\sqrt{n}\}. \end{aligned} \tag{2}$$

The average number of articles tested before an alarm is raised at the upper limit is then $A(n)=n/P$, where

$$P = \frac{1}{\sqrt{2\pi}} \int_{B-k\sqrt{n}}^{\infty} \exp\left(-\tfrac{1}{2}z^2\right) dz. \tag{3}$$

For any given n and B, the average number of articles tested (or "average amount of inspection") $A(n)$ is a function of k, whose values can be calculated with the aid of a set of normal tables.

In Charts I and II, $A(n)$ is plotted against k for various values of n, while $a=2000$ and $a=5000$, respectively. Both charts show clearly the superiority of large sample sizes over a wide range of k values. For instance, $n=10$ is more powerful than $n=5$ for any value of k less than 1.2 or 1.3, and $n=20$ is more powerful than $n=5$ for k less than 1.0 or 1.1. In particular, for $a=5000$ and $k=0.7$, the average amount of inspection is 80 for $n=5$, 40 for $n=10$, and only 30 for $n=20$. The saving of inspection is even greater when samples of 50 are used, but the range of k for which such a large sample size is powerful is rather restricted ($k<0.5$, say). Only for the detection of very large shifts of the mean will small samples be more powerful than large samples.

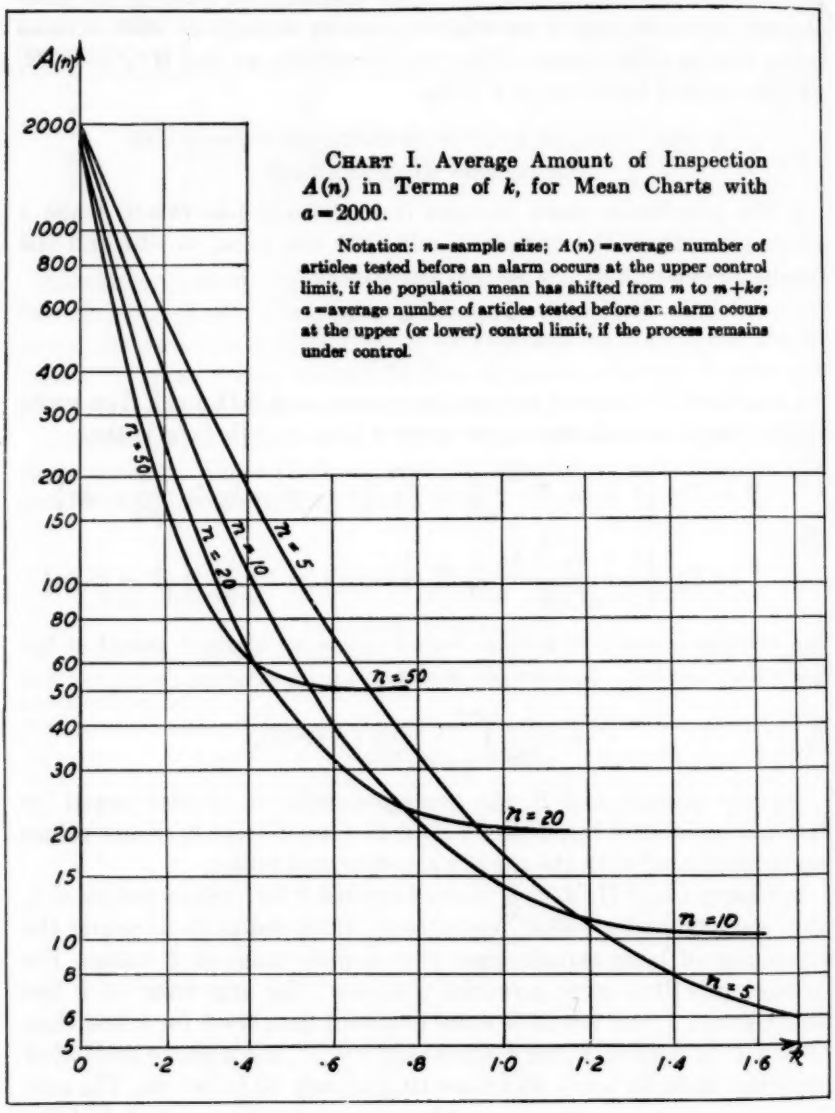


CHART I. Average Amount of Inspection $A(n)$ in Terms of k, for Mean Charts with $a = 2000$.

Notation: n = sample size; $A(n)$ = average number of articles tested before an alarm occurs at the upper control limit, if the population mean has shifted from m to $m + k\sigma$; a = average number of articles tested before an alarm occurs at the upper (or lower) control limit, if the process remains under control.

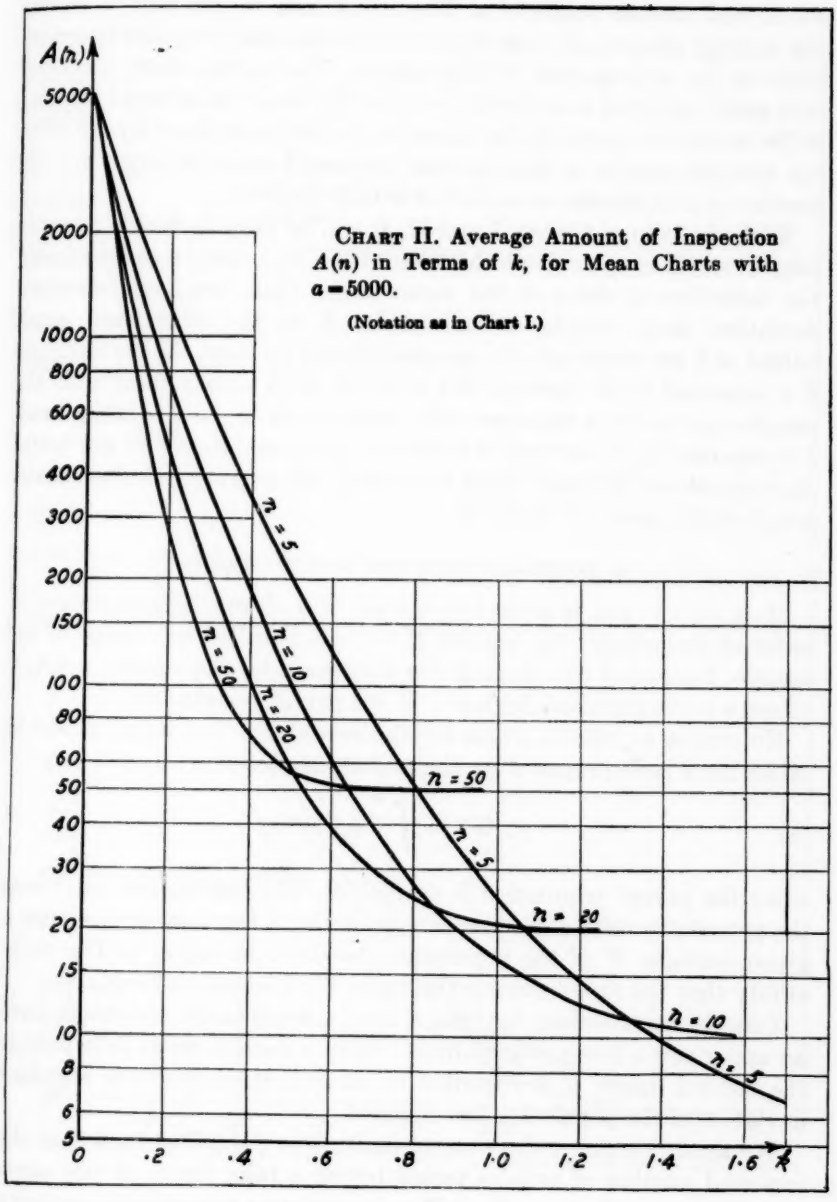


CHART II. Average Amount of Inspection $A(n)$ in Terms of k, for Mean Charts with $a = 5000$.

(Notation as in Chart I.)

In Chart III, $A(n)$ is plotted against k for the fixed sample size $n=10$ and various values of a. They show that we are able to reduce the average amount of inspection by making a smaller, that is, by increasing the average rate of false alarms. The curves show, however, that small values of a are useful only for the detection of small changes of the population mean. If, for instance, $a=5000$ is replaced by $a=1000$, the average number of false alarms becomes 5 times as large, but the number of real alarms when $k=0.6$ is only doubled.

With the help of Charts I and II, it will be easy to decide whether large or small samples should be taken. If we are mainly concerned with the detection of shifts of the mean larger than (say) one standard deviation, small samples are advisable. If, on the other hand, small values of k are expected, the samples should be large. If, for instance, k is expected to lie between 0.4 and 0.8, both charts show that the sample size $n=20$ is more powerful than $n=10$. If, on the other hand, k is expected to lie between 0.8 and 1.2 (say), samples of 10 are better than samples of 20. Once a and n are fixed, the control limits can be determined by means of Table II.

4. CONTROL LIMITS FOR RANGE CHARTS

If $x_1, x_2, \cdots, x_n$ is a random sample of n observations arranged in order of magnitude, the variate $R=x_n-x_1$ is called the range of the sample. Instead of the range R, we shall consider the variate $w=R/\sigma$, where σ is the standard deviation of the parent population.

No simple expression exists for the probability law $\phi_n(w)$ of w, but tables have been prepared for the probability integral

$$p_n(W) = \int_0^W \phi_n(w)dw \tag{4}$$

when the parent population is normal [1]. This expression represents the probability that a random sample of size n has a range less than a given multiple W of the population standard deviation σ. The probability that the range exceeds the value $W\sigma$ is equal to $1-p_n(W)$.

Like the control chart for means, most control charts for ranges have an upper and a lower control limit. When a sample range falls outside the control limits, it is regarded as an indication that the standard deviation of the population has changed.

We shall determine the control limits $W_1\sigma \leqq R \leqq W_2\sigma$ such that the expected number of articles tested before a false alarm at the upper limit $W_2\sigma$, or at the lower limit $W_1\sigma$, is equal to a preassigned number a for any sample size n.

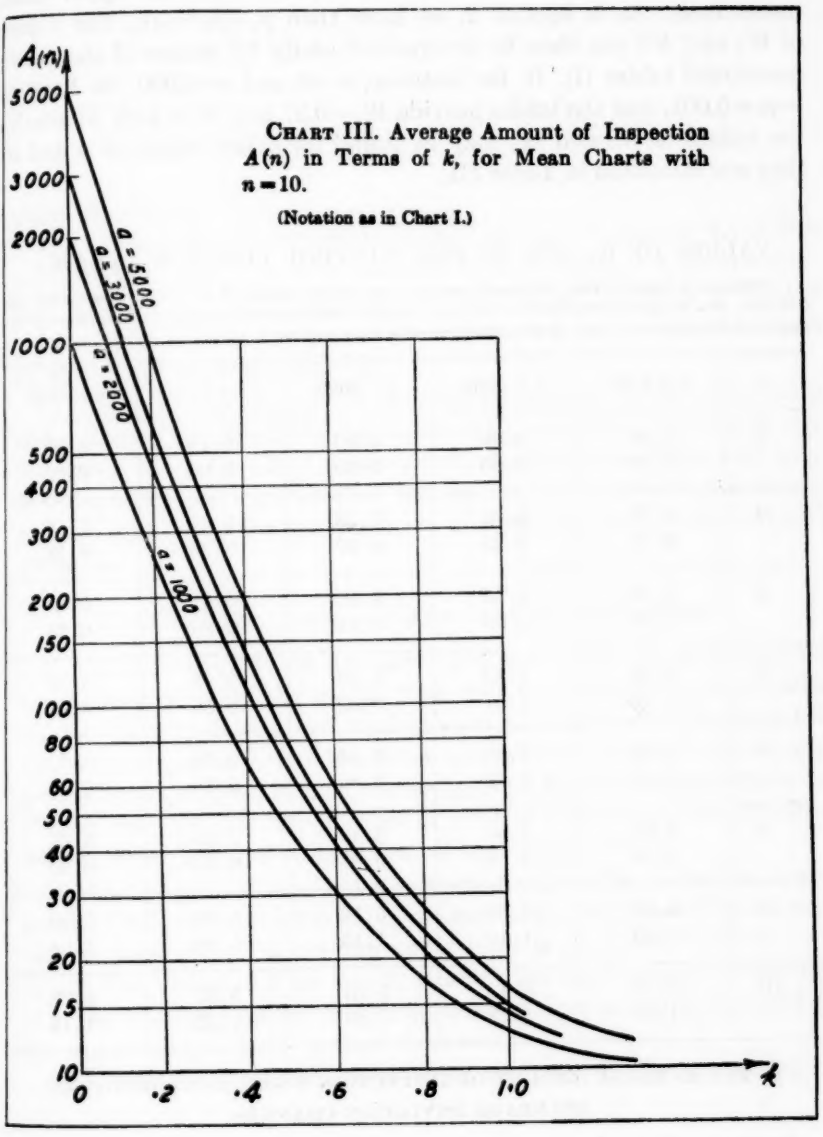


CHART III. Average Amount of Inspection $A(n)$ in Terms of k, for Mean Charts with $n = 10$.

(Notation as in Chart I.)

Let $p_1 = p_n(W_1)$ and $p_2 = p_n(W_2)$ be the probabilities that the range of a random sample falls below the lower and above the upper limit, respectively. As in Section 2, we have then $p_1 = p_2 = n/a$. The values of W_1 and W_2 can then be determined easily by means of the above mentioned tables [1]. If, for instance, $n = 5$ and $a = 5000$, we have $p_1 = p_2 = 0.001$, and the tables provide $W_1 = 0.37$ and $W_2 = 5.48$. Similarly, the values of W_1 and W_2 may be found for other values of n and a; they are tabulated in Table III.

TABLE III

VALUES OF W_1 AND W_2 FOR CONTROL LIMITS $W_1\sigma \leqq R \leqq W_2\sigma$

Notation: n = sample size; R = sample range; σ = population standard deviation; $W_1\sigma$ = lower control limit; $W_2\sigma$ = upper control limit; a = average number of articles tested before an alarm occurs at the upper (or lower) control limit, if the process remains under control.

n	$a = 1000$	$a = 2000$	$a = 3000$	$a = 4000$	$a = 5000$
3	4.64 0.10	4.00 0.06	5.06 0.06	5.18 0.05	5.25 0.04
4	4.78 0.31	5.05 0.25	5.23 0.22	5.31 0.20	5.40 0.18
5	4.88 0.55	5.15 0.46	5.30 0.42	5.41 0.39	5.48 0.37
6	4.96 0.78	5.23 0.67	5.38 0.62	5.48 0.58	5.55 0.56
7	5.02 0.93	5.29 0.86	5.44 0.80	5.54 0.76	5.63 0.73
8	5.07 1.16	5.34 1.04	5.47 0.97	5.58 0.93	5.67 0.90
9	5.12 1.33	5.38 1.19	5.53 1.13	5.64 1.09	5.70 1.05
10	5.16 1.47	5.42 1.33	5.57 1.27	5.67 1.22	5.73 1.18

5. THE AVERAGE AMOUNT OF INSPECTION WHEN THE POPULATION STANDARD DEVIATION CHANGES

Suppose that the standard deviation of the parent population changes from σ to $\sigma' = k\sigma$, $k > 1$. The variate $w' = R/\sigma'$ has then the same distribution as previously the variate $w = R/\sigma$. It follows that

the probability that the range R of a random sample falls above the upper limit $W_2\sigma$ is

$$P = 1 - \Pr\{R < W_2\sigma\} = 1 - \Pr\left\{w' < W_2\frac{\sigma}{\sigma'}\right\}$$

$$= 1 - \Pr\left\{w' < \frac{W_2}{k}\right\} = 1 - \int_0^W \phi_n(w)dw, \tag{5}$$

where $W = W_2/k$.

The average number of articles tested before a change of the standard deviation from σ to $k\sigma$ is detected is, as in Section 3, $A(n) = n/P$. $A(n)$ depends on n and k and also on the number a defined in Section 4. Its relation to k for any given values of a and n can be easily obtained by means of the range tables [1].

In Charts IV and V, $A(n)$ is plotted against k for some values of n and a. Chart IV shows that little is gained by increasing the sample size from $n=5$ to $n=10$. Chart 5 shows that the amount of inspection may be reduced by making a smaller, but, as for mean charts, small values of a are useful only for the detection of small changes in the standard deviation.

6. CONTROL LIMITS FOR MEAN CHARTS USING RUNS

It was shown in [3] that the power of mean charts for small samples can be improved by the use of runs. For such charts, control limits $m \pm B\sigma/\sqrt{n}$ are determined, and as soon as λ successive $\bar{x}$ values fall above the upper or below the lower control limit, alarm is raised and production is stopped to allow investigation. We shall again determine B such that the average number of articles tested between two successive false alarms is independent of the sample size.

If P is the probability that a random $\bar{x}$ value falls above the upper control limit, the average number of samples that will pass before λ successive $\bar{x}$ values fall above the upper limit is [3] $S = P^{-1} + P^{-2} + \cdots + P^{-\lambda}$. The values of P for any B, k, and n are again given by equation (3). The average number of articles tested before an alarm is raised at the upper control limit, is then $A(n) = nS$.

In particular, a *false* alarm is raised when $k=0$. The probability P then becomes

$$p = \frac{1}{\sqrt{2\pi}}\int_B^\infty \exp\left(-\tfrac{1}{2}z^2\right)dz, \tag{6}$$

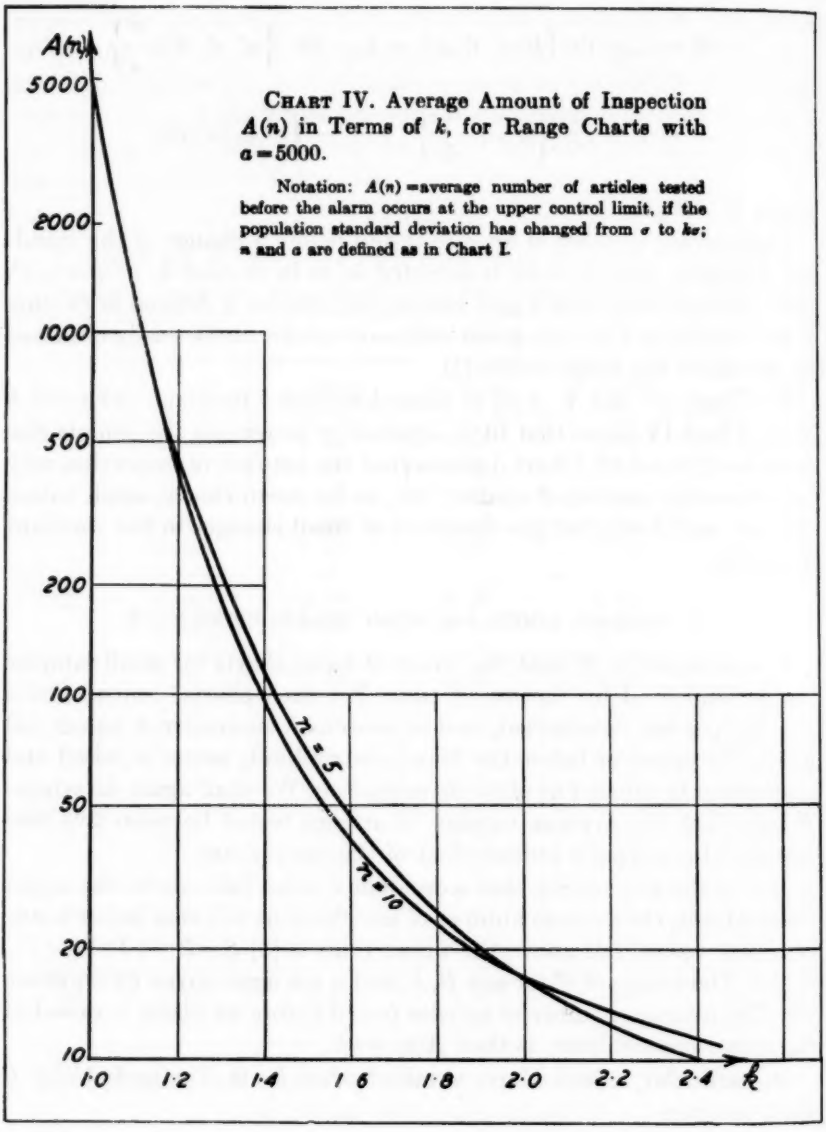


CHART IV. Average Amount of Inspection $A(n)$ in Terms of k, for Range Charts with $a = 5000$.

Notation: $A(n)$ = average number of articles tested before the alarm occurs at the upper control limit, if the population standard deviation has changed from σ to $k\sigma$; n and a are defined as in Chart I.

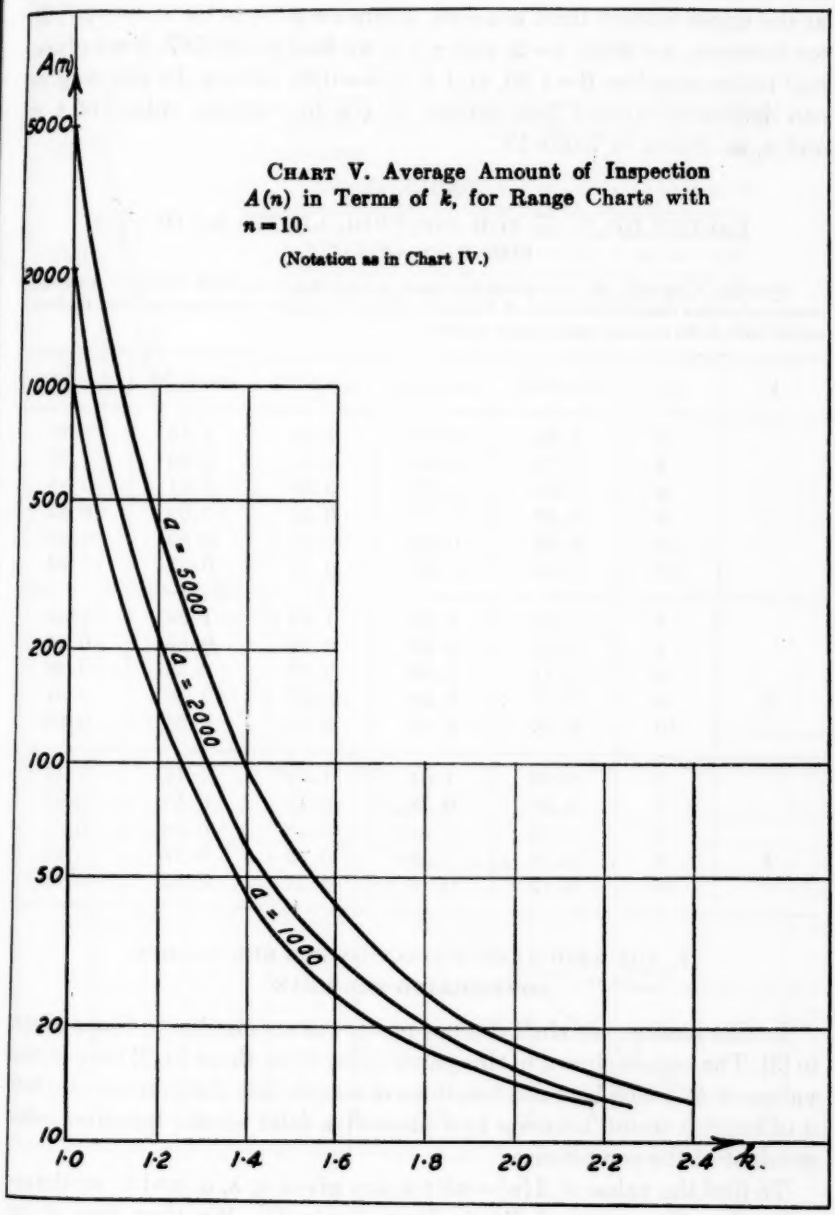


CHART V. Average Amount of Inspection $A(n)$ in Terms of k, for Range Charts with $n = 10$.

(Notation as in Chart IV.)

and the average number of articles tested before a false alarm is raised at the upper control limit is $a=ns$, where $s=p^{-1}+p^{-2}+\cdots+p^{-\lambda}$. If, for instance, $a=5000$, $\lambda=2$, and $n=4$, we find $p=0.0287$. A set of normal tables supplies $B=1.90$, and $B/\sqrt{n}=0.95$ follows. In this way we can determine control limit factors $B/\sqrt{n}$ for various values of λ, a, and n, as shown in Table IV.

TABLE IV

VALUES OF $B/\sqrt{n}$ FOR CONTROL LIMITS $m\pm(B/\sqrt{n})\sigma$ FOR RUN CHARTS

Notation: n = sample size; m = population mean; σ = population standard deviation; a = average number of articles tested before a run of λ sample means occurs above the upper (or below the lower) control limit, if the process remains under control.

λ	n	$a=1000$	$a=2000$	$a=3000$	$a=4000$	$a=5000$
2	1	1.85	2.00	2.09	2.15	2.20
	4	0.76	0.85	0.90	0.93	0.95
	5	0.65	0.73	0.78	0.81	0.83
	8	0.47	0.53	0.57	0.60	0.62
	10	0.40	0.46	0.49	0.52	0.53
	20	0.23	0.28	0.31	0.33	0.34
3	1	1.26	1.39	1.47	1.52	1.56
	4	0.49	0.56	0.59	0.63	0.65
	5	0.41	0.48	0.52	0.55	0.56
	8	0.33	0.34	0.37	0.40	0.41
	10	0.23	0.29	0.32	0.34	0.35
4	1	0.89	1.01	1.08	1.13	1.16
	4	0.31	0.38	0.42	0.45	0.47
	5	0.25	0.32	0.35	0.38	0.40
	8	0.16	0.22	0.25	0.27	0.28
	10	0.12	0.18	0.21	0.22	0.24

7. THE AMOUNT OF INSPECTION FOR RUN CHARTS CONTROLLING THE MEAN

In this section, we shall discuss power curves similar to those given in [3]. The curves shown in this paper differ from those in [3] only in the values of B, which here are functions of n such that the average number a of articles tested between two successive false alarms becomes independent of the sample size.

To find the value of $A(n)=nS$ for any given n, λ, a, and k, we determine first the value of $B/\sqrt{n}$ from Table IV. We then find P by equation (3), using a set of normal tables, and deduce $S=P^{-1}+P^{-2}$

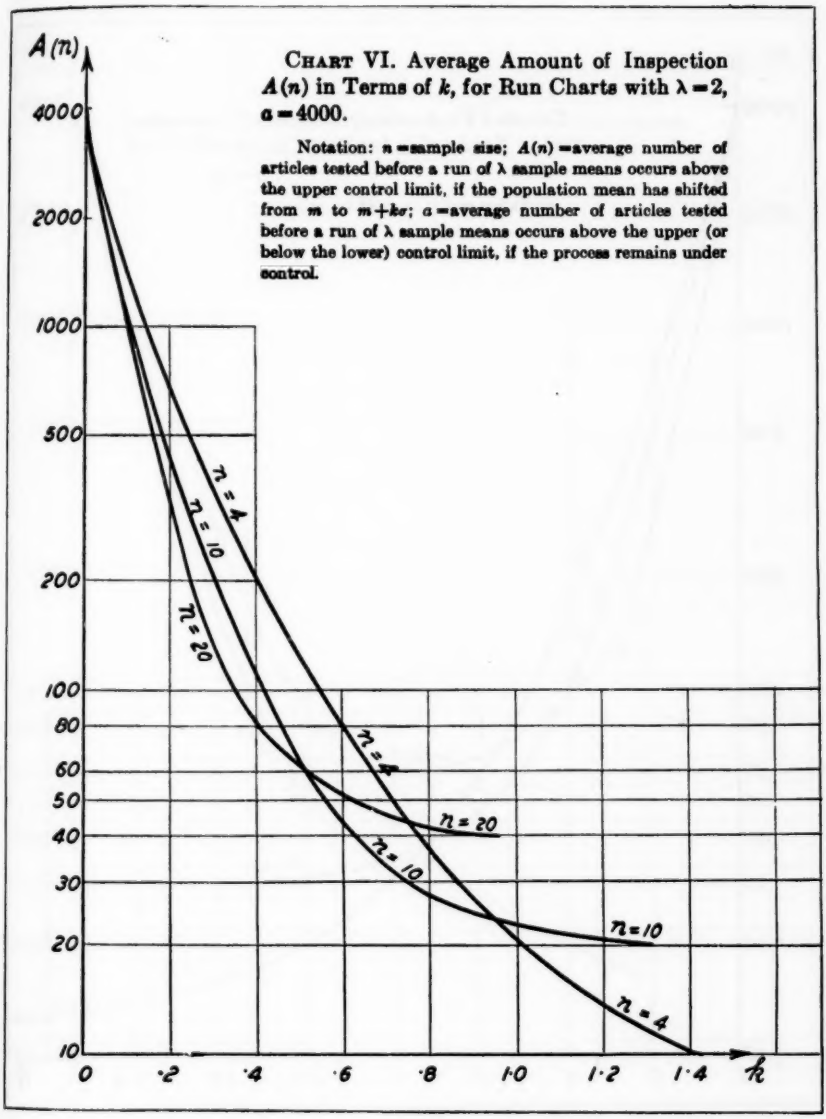


CHART VI. Average Amount of Inspection $A(n)$ in Terms of k, for Run Charts with $\lambda = 2$, $a = 4000$.

Notation: n = sample size; $A(n)$ = average number of articles tested before a run of λ sample means occurs above the upper control limit, if the population mean has shifted from m to $m+k\sigma$; a = average number of articles tested before a run of λ sample means occurs above the upper (or below the lower) control limit, if the process remains under control.

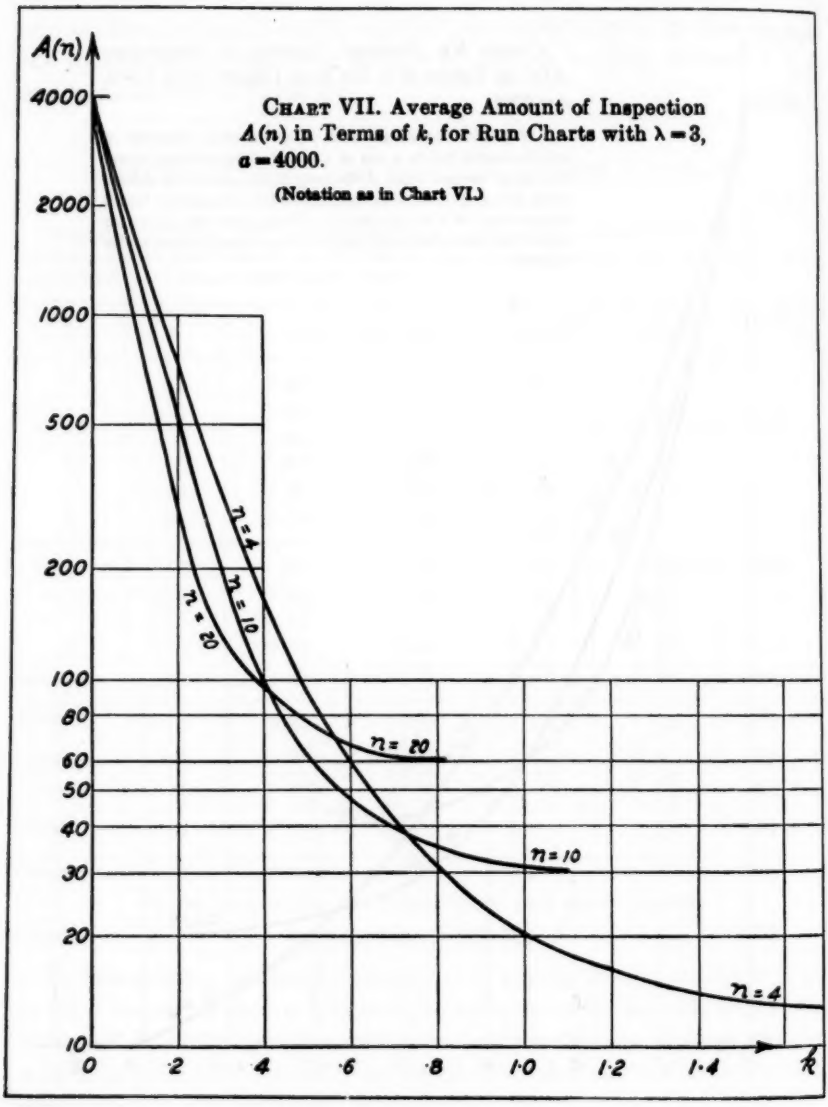


CHART VII. Average Amount of Inspection $A(n)$ in Terms of k, for Run Charts with $\lambda = 3$, $a = 4000$.

(Notation as in Chart VI.)

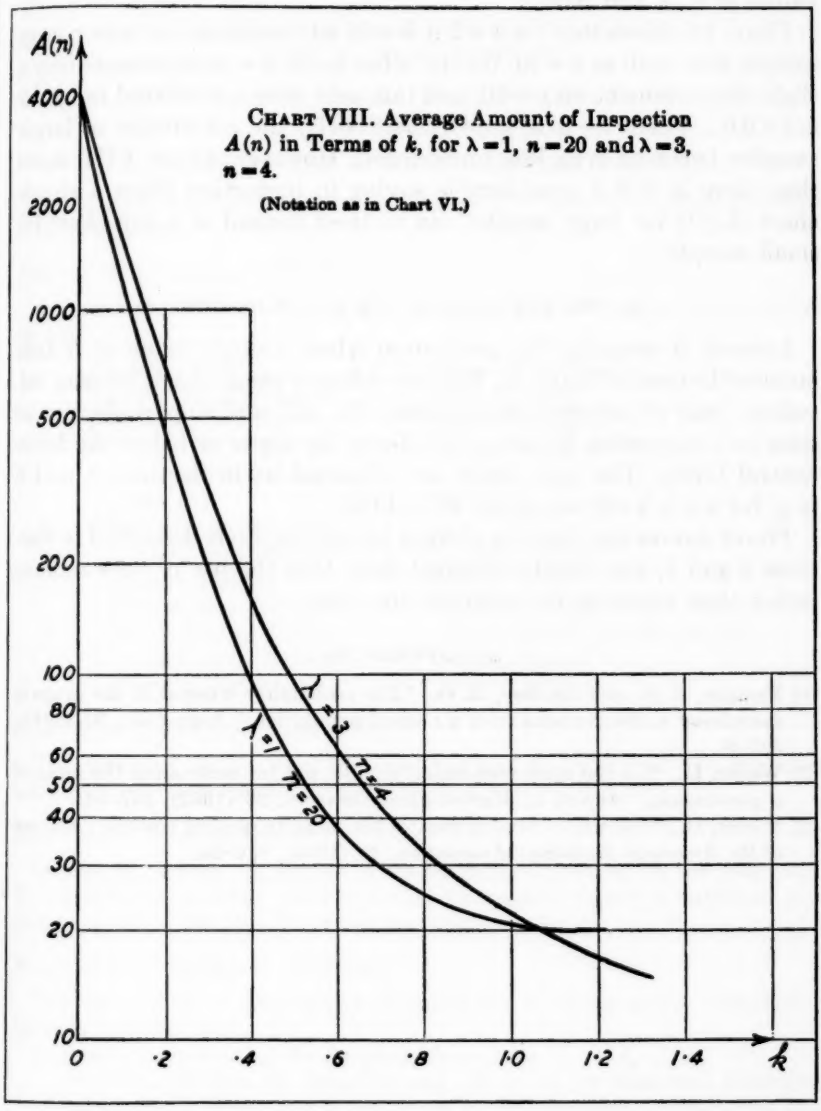


CHART VIII. Average Amount of Inspection $A(n)$ in Terms of k, for $\lambda = 1$, $n = 20$ and $\lambda = 3$, $n = 4$.

(Notation as in Chart VI.)

$+ \cdots + P^{-\lambda}$. (Actually, to plot the curves, we found it more convenient to start with given values of P and to deduce corresponding values of k, S, and $A(n)$.)

Chart VI shows that for $\lambda = 2$ it is still advantageous to take a large sample size, such as $n = 10$. On the other hand, $n = 20$ represents only a slight improvement on $n = 10$, and this only over a restricted range for $k (k < 0.6)$. When $\lambda = 3$ is used (Chart VII), the advantage of larger samples becomes even less pronounced. However, Chart VIII shows that there is still a considerable saving in inspection when a simple chart ($\lambda = 1$) for large samples can be used instead of a run chart for small samples.

8. THE USE OF RUNS FOR RANGE CHARTS

Instead of stopping the production when a single value of R falls outside the control limits R_1, R_2 of an ordinary range chart, we may calculate a pair of narrower control limits R_1', R_2', and stop production as soon as λ successive R values fall above the upper or below the lower control limits. The new limits are obtained as in Sections 4 and 6 (e.g. for $n = 5$, $\lambda = 2$, we obtain $R_2' = 4.08\sigma$).

Power curves can then be plotted by the methods described in Sections 5 and 7, but results obtained show that the use of runs reduces rather than improves the power of the chart.

9. REFERENCES

[1] Pearson, E. S., and Hartley, H. O., "The probability integral of the range in samples of n observations from a normal population," *Biometrika*, 32 (1942), 301–8.

[2] Weiler, H., "On the most economical sample size for controlling the mean of a population," *Annals of Mathematical Statistics*, 23 (1952), 247–54.

[3] Weiler, H., "The use of runs to control the mean in quality control," *Journal of the American Statistical Association*, 48 (1953), 816–25.

PROCEEDINGS

AMERICAN STATISTICAL ASSOCIATION
113TH ANNUAL MEETING

SHOREHAM HOTEL, WASHINGTON, D. C.
DECEMBER 28, 1953

MINUTES OF THE ANNUAL BUSINESS MEETING

The meeting was called to order by William G. Cochran, outgoing President of the Association.

Report of the Committee on Elections

A Report of the Committee on Elections shows the following officers elected for 1954:

President Elect	Ralph J. Watkins
Vice President (1954–56)	Henry Scheffé
Directors (1954–56)	Jacob Marschak
	Donald C. Riley
Representative at Large (1954–55)	Daniel B. deLury
District Representatives	
Northeastern District	Chester I. Bliss
Eastern District	George Garvy
Southeastern District	Frank A. Hanna
North Central District	Paul R. Rider, Lucile Derrick
South Central District	William Kester
Western District	John C. McKee

Report of the Board of Directors for 1953

The Report of the Board of Directors was read and accepted. The Report is published separately following the Minutes of this Meeting.

Report of the Secretary-Treasurer for 1953

Samuel Weiss read the Report of the Secretary-Treasurer for the year 1953. The Report was accepted. The Secretary-Treasurer's Report is published separately following the Minutes of this Meeting.

Report of the Committee on Resolutions

The Committee on Resolutions presented the following to the membership for their consideration:

1. Resolution regarding the Program Committee
   RESOLVED that the members and officers of the American Statistical Association express deep appreciation for the excellent program prepared by members of the Program Committee under the leadership of Herbert Solomon, Chairman.
2. Resolution regarding the Local Arrangements Committee and the Washington, D. C. Chapter.

RESOLVED that the members and officers of the American Statistical Association express their profound appreciation to the Local Arrangements Committee under the Chairmanship of Donald C. Riley and to all of the individuals of the Washington, D. C. Chapter for their outstanding work and hospitality in connection with the arrangements for the 113th Annual Meeting of the Association.

The resolutions were approved.

Report on the Schedule of Forthcoming Meetings

The forthcoming meetings of the Association are scheduled as follows:

1954—Annual Meeting—Montreal, Canada—September 10–13, 1954
1954—Regional Meeting—San Francisco Regional Conference—Berkeley, California—December 1954
1955—Annual Meeting—New York City—December 27–29, 1955

There being no new business, the meeting was adjourned.

REPORT OF THE BOARD OF DIRECTORS, 1953

Activities of the Association during 1953 were vigorous and widespread in scope. This was coupled with a most successful financial year. Details of finance and membership will be found in the Secretary-Treasurer's Report.

1. *Sections and Committees*

The Board has granted sectional status to the Social Statistics Section, formerly the Committee on Statistics in the Social Sciences. This Committee requested sectional status at the Meeting of the Incoming Board and Council in December, 1952. The matter was referred to the Committee on Committees, which, after reviewing the final charter of the proposed section, recommended to the Board that approval be given. The approval of this Charter makes this the fourth Section of the Association; the other three being the Biometric Section, the Business and Economics Statistics Section and the Section on the Training in Statistics.

These four Sections and the Committee on Statistics in the Physical Sciences have been extremely active in the formulation and planning of the Annual Meeting programs, as well as the planning for a number of successful Regional Meetings. Without any question, the formation of the Sections has been extremely beneficial to the Association. The variety of interests of the membership of the Association is best served by the activity of these Sections.

Two *ad hoc* committees were constituted by the Board this year. They are the *Ad Hoc* Committee on Publications Policy and the *Ad Hoc* Committee on Statistical Standards and Organization. Both groups were asked to make their Reports to the Board so that action may be taken on their recommendations in these two important areas.

The Board wishes to commend the 1953 Program and Local Arrangement Committees. The former, with Herbert Solomon as Chairman, has done an outstanding job of presenting a well-balanced, varied group of sessions for the Annual Meeting. The Local Arrangement Committee, chaired by Donald Riley, has worked very hard to ensure a successful meeting. Their efforts for ASA, in cooperation with the other societies meeting jointly, have produced a memorable convention.

2. *Abstracts*

This year for the first time the Association has published abstracts of papers presented at the previous Annual Meeting in the *Journal of the American Statistical Association.* The editor of these abstracts was Arman Alchian. The Board is very pleased to announce that Professor Alchian has agreed to continue in this position for the coming year. It is hoped that, with cooperation from persons presenting papers at this Annual Meeting, the abstracts will appear in an early issue of the 1954 volume of the *Journal.*

3. *New Constitution*

The new Constitution, approved by the membership in a mail ballot in 1953, will go into effect January 1, 1954. The changes incorporated into the new version are, for the most part, procedural rather than substantive, and should facilitate the operation of the Association. The draft version, which was ratified by the membership, was printed in the June–July, 1952 issue of the *American Statistician.*

4. *New Chapters*

During 1953 four new Chapters were granted Charters by the Board of Directors. The new groups are located in Tulsa, New Orleans, Puerto Rico and Milwaukee. The Board extends a welcome to these new Chapters.

The total number of active Chapters has now reached 31. Chapter meetings, often in cooperation with local chapters of other societies, offer a wide variety of topics to the membership in the fields of statistical interest.

5. *Conferences and Meetings*

During 1953 the Business and Economic Statistics Section sponsored two regional conferences. The first was held on April 30 and May 1 in cooperation with the Graduate School of Industrial Administration at the Carnegie Institute of Technology and the Pittsburgh Chapter of the American Marketing Association and the American Society for Quality Control. The Conference was devoted to modern statistical methods in business and industry.

The second Conference, sponsored jointly with the Wharton School of the University of Pennsylvania, within its theme of business statistics was divided into three main sessions—on capital outlays, production scheduling and sales forecasting. This Conference took place in Philadelphia on June 11 and 12. Both Conferences were well attended and very successful.

6. *New Monograph*

Following the Association's policy of a more vigorous program of publication, the Board has voted to publish as a monograph the complete Report of the ASA Committee to Advise the National Research Council Committee for Research in the Problems of Sex. This Report, dealing with the volume on the human male entitled "Sexual Behavior in the Human Male," by Professor Kinsey and his associates, will be available in 1954. Certain sections of the Report are also being published as articles in the *Journal of the American Statistical Association.*

7. *New Appointments of ASA Representatives*

Samuel S. Wilks was elected by the Council to continue to serve as the ASA Representative to the Social Science Research Council for a three-year term.

The National Bureau of Economic Research announced the resignation, due to illness, of Frederick C. Mills as ASA Representative to its Board of Directors. The Board has appointed W. Allen Wallis to complete the unexpired portion of Dr. Mills' term. A resolution was received from the National Bureau of Economic Research which expressed deep regret for the necessity of Dr. Mills' resignation and thanking the ASA for his valuable services during his tenure of office.

8. *Journal of the American Statistical Association*

The Board approved an increase in the budget for the *Journal* for 1953, with the result that more articles were published in the 1953 volume, and the number of pages increased by approximately 25 per cent over 1952. In addition, the Board approved an increase in the funds for editorial assistance, which heretofore has been supplied by the University of Chicago.

9. *Reduction in Dues for Foreign Members*

The Board voted to reduce the dues for persons residing outside North America, beginning in 1954. The reduction from $8.00 to $5.00 yearly will make it easier for many foreign members in terms of dollar exchange. It is expected that the slight loss in income from this reduction will be more than offset by the increase in new members outside North America. A 1954 drive for new foreign members is planned.

10. *Future Meetings*

Future plans of the Association call for the 1954 Annual Meeting to be held in Montreal, Canada, on September 10–13. Two Regional Meetings are also being planned for 1954. The Chicago Chapter, in cooperation with other Chapters in its District, has begun work on a meeting to be held in the spring, probably during April. A Western Regional Meeting has been scheduled for San Francisco in December of 1954, in conjunction with the American Association for the Advancement of Science. Final dates and headquarters for this conference have not yet been chosen. Advance publicity will be issued for both of these Regional Meetings.

The Montreal Meeting was planned for September to leave Christmas week free for members who usually attend annual Meetings, and to attract those who do not attend late December Meetings. In this sense it is an experiment designed to guide the Association in the selection of other meeting dates. The widest understanding of the nature of the experiment will help its success.

REPORT OF THE SECRETARY-TREASURER, 1953

A drive for an increase in new members which started in 1952 continued actively in 1953. This, together with a policy of careful economy, has resulted in increasing the Association's surplus by a wider margin than ever before.

Income for 1953 was budgeted at $52,650. The actual total 1953 income is $60,377.31. This difference is due to a rise in income of most budgeted items, but primarily results from increases in receipts from membership dues, sales of publications and subscriptions. Expenses were budgeted at $50,612. The actual total 1953 figure is $50,433.83. All expenses have been kept close to budget level, and small savings have been made on a number of different items. Thus,

the Association shows an increase to surplus of $9,943.48 at the end of 1953. This brings the total surplus to over $26,000, about halfway to the goal of a surplus equal to one year's income.

For the second successive year the number of members has shown a significant increase. At the beginning of 1953 the membership totaled 4,655. The number of new members for 1953 was 639, and 25 others reinstated their membership. At the end of 1953 approximately 400 members have been dropped from the rolls because of resignation, death or non-payment of dues. Thus, the net membership growth for 1953 is 264, and the Association starts 1954 with a total of more than 4,900 members. This is the highest level of membership ever reached by the Association. It is expected that the drive for more foreign members, as mentioned in the Board of Directors' Report, will add substantially to the membership in 1954.

Subscriptions to the *Journal of the American Statistical Association* have also been increasing. At the end of 1952 there were 1,248 subscribers to the *Journal*, while at the end of 1953 the figure had risen to 1,356. This increase is expected to continue in 1954.

Financial Recommendations

The Report of the Treasurer, shown separately, emphasizes that 1953 was the fourth year in succession in which the Association has accrued surplus. The Board of Directors for the past two years has recommended that the surplus be increased until it equals the income of the Association for one year. At that time it is felt that the Association will be in a much stronger position to expand its activities on a much wider scale. With this in mind, the Treasurer has planned to budget between $2,000 and $3,000 per year for addition to surplus until this goal is reached. The proposed income for 1954 is budgeted at approximately $55,000, while expense has been calculated at $52,500, leaving $2,500, for addition to surplus. Income has been figured very conservatively, while expense has been approximated as closely as possible, with the expectation that the surplus may be somewhat larger than budgeted.

April 19, 1954

To the Board of Directors of
American Statistical Association.

I have examined the attached financial statements of American Statistical Association relating to the year ended December 31, 1953. My examination was made in accordance with generally accepted auditing standards and, accordingly, included such tests of the accounting records and such other auditing procedures as were considered necessary in the circumstances.

The recorded cash receipts for the year were traced in the deposits shown on the bank statements and the amounts for dues and subscriptions were tested with the membership and subscription records. The paid checks were inspected and related vouchers tested in support of cash disbursements for the year. The bank balances were reconciled with amounts reported directly to me by the depositaries and the cash on hand at December 31, 1953 was verified by inspection. I did not check the membership and subscription records in detail or make any independent verification of the inventory of old *Journals*, the office records of which are based, in part, on data assembled in prior years.

In accordance with a resolution of the Board of Directors, the expense incurred

in publishing a directory, distributed to the membership in 1951, is being spread over a three-year period although such costs would appear to be applicable primarily to the year 1951. The accounts for the year ended December 31, 1953 reflect a charge of $831.86, representing the allocated portion of the directory expense applicable to that period.

In my opinion, the accompanying statements present fairly the position of American Statistical Association at December 31, 1953, and the results of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent, except as mentioned in the preceding paragraph, with that of the preceding year.

JAMES G. JESTER

AMERICAN STATISTICAL ASSOCIATION
BALANCE SHEET

Assets

	December 31, 1953	1952
Cash in banks and on hand	$52,431.05	$39,503.14
Accounts receivable	2,783.14	1,099.58
Investment in United States Savings Bonds, Series G, due 1962, at cost	3,100.00	3,100.00
Inventory of old *Journals*, at approximate cost	2,137.69	1,909.85
Inventory of Monograph on Acceptance Sampling, at cost	129.24	233.93
Inventory of Emblems, at cost	415.50	463.50
Furniture and fixtures, at cost less depreciation	2,088.78	2,192.42
Deferred Charges:		
Deferred Membership Directory expense		831.86
Other	945.55	1,007.87
	$64,030.95	$50,342.15

Liabilities and Net Worth

Accounts payable	$10,430.60	$ 5,930.38
Deferred income (collections applicable to subsequent years)		
Dues	$16,827.00	$16,179.00
Subscriptions	5,817.77	5,524.48
Other	466.84	2,865.16
	$23,111.61	$24,568.64
Net Worth:		
Life Membership reserve	$ 3,579.92	$ 2,877.79
Surplus, per statement	26,908.82	16,965.34
	$30,488.74	$19,843.13
	$64,030.95	$50,342.15

AMERICAN STATISTICAL ASSOCIATION
STATEMENT OF INCOME AND SURPLUS ACCOUNTS

	Year ended December 31, 1953	1952
Income:		
Dues—Current year	$38,607.00	$37,101.00
—Prior year	852.00	184.00
Life membership income	(90.13)	166.31
Subscriptions—*Journal*	10,134.80	9,543.00
—*American Statistician*	443.68	428.35
Advertising—*Journal*	1,415.99	1,376.75
—*American Statistician*	263.97	217.72
Sales—*Journal*	1,937.13	1,180.84
—*American Statistician*	148.02	82.65
—Acceptance Sampling	302.05	243.00
—Emblems, less cost of sales	83.28	11.00
—Membership Directory	22.50	45.00
—*Biometrics*	568.37	453.25
—Other	45.00	35.75
Mailing list income	911.11	704.17
Interest income	911.84	541.37
Annual meeting—see note	1,559.52	
Reimbursement of overhead expenses:		
Bureau of Mines Project	2,207.06	2,000.00
Miscellaneous	54.12	67.33
	$60,377.31	$54,381.49
Expense:		
Salaries	$12,260.07	$14,716.81
Publications—Schedule I	24,968.23	20,417.77
Promotion	861.89	702.30
Rent	2,400.00	2,400.00
Travel and secretarial expense	800.83	1,426.17
Supplies	2,240.92	2,434.54
Postage	1,849.26	1,163.45
Telephone and telegraph	729.02	563.71
Accounting services	970.00	970.00
Committee expense	1,269.75	575.50
Annual meeting expense	645.99	550.72
Miscellaneous expenses—Schedule I	1,437.87	1,928.83
	$50,433.83	$47,849.80
Excess of income over expense for the year	$ 9,943.48	$ 6,531.69
Add: Surplus account, at beginning of year	16,965.34	10,433.65
Surplus account, at end of year	$26,908.82	$16,965.34

Note: Includes $303.29 relating to receipts from 1952 meeting.

in publishing a directory, distributed to the membership in 1951, is being spread over a three-year period although such costs would appear to be applicable primarily to the year 1951. The accounts for the year ended December 31, 1953 reflect a charge of $831.86, representing the allocated portion of the directory expense applicable to that period.

In my opinion, the accompanying statements present fairly the position of American Statistical Association at December 31, 1953, and the results of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent, except as mentioned in the preceding paragraph, with that of the preceding year.

JAMES G. JESTER

AMERICAN STATISTICAL ASSOCIATION
BALANCE SHEET

Assets

	December 31, 1953	1952
Cash in banks and on hand	$52,431.05	$39,503.14
Accounts receivable	2,783.14	1,099.58
Investment in United States Savings Bonds, Series G, due 1962, at cost	3,100.00	3,100.00
Inventory of old *Journals*, at approximate cost	2,137.69	1,909.85
Inventory of Monograph on Acceptance Sampling, at cost	129.24	233.93
Inventory of Emblems, at cost	415.50	463.50
Furniture and fixtures, at cost less depreciation	2,088.78	2,192.42
Deferred Charges:		
Deferred Membership Directory expense		831.86
Other	945.55	1,007.87
	$64,030.95	$50,342.15

Liabilities and Net Worth

Accounts payable	$10,430.60	$ 5,930.38
Deferred income (collections applicable to subsequent years)		
Dues	$16,827.00	$16,179.00
Subscriptions	5,817.77	5,524.48
Other	466.84	2,865.16
	$23,111.61	$24,568.64
Net Worth:		
Life Membership reserve	$ 3,579.92	$ 2,877.79
Surplus, per statement	26,908.82	16,965.34
	$30,488.74	$19,843.13
	$64,030.95	$50,342.15

American Statistical Association
Statement of Income and Surplus Accounts

	Year ended December 31, 1953	1952
Income:		
Dues—Current year	$38,607.00	$37,101.00
—Prior year	852.00	184.00
Life membership income	(90.13)	166.31
Subscriptions—*Journal*	10,134.80	9,543.00
—*American Statistician*	443.68	428.35
Advertising—*Journal*	1,415.99	1,376.75
—*American Statistician*	263.97	217.72
Sales—*Journal*	1,937.13	1,180.84
—*American Statistician*	148.02	82.65
—Acceptance Sampling	302.05	243.00
—Emblems, less cost of sales	83.28	11.00
—Membership Directory	22.50	45.00
—*Biometrics*	568.37	453.25
—Other	45.00	35.75
Mailing list income	911.11	704.17
Interest income	911.84	541.37
Annual meeting—see note	1,559.52	
Reimbursement of overhead expenses:		
Bureau of Mines Project	2,207.06	2,000.00
Miscellaneous	54.12	67.33
	$60,377.31	$54,381.49
Expense:		
Salaries	$12,260.07	$14,716.81
Publications—Schedule I	24,968.23	20,417.77
Promotion	861.89	702.30
Rent	2,400.00	2,400.00
Travel and secretarial expense	800.83	1,426.17
Supplies	2,240.92	2,434.54
Postage	1,849.26	1,163.45
Telephone and telegraph	729.02	563.71
Accounting services	970.00	970.00
Committee expense	1,269.75	575.50
Annual meeting expense	645.99	550.72
Miscellaneous expenses—Schedule I	1,437.87	1,928.83
	$50,433.83	$47,849.80
Excess of income over expense for the year	$ 9,943.48	$ 6,531.69
Add: Surplus account, at beginning of year	16,965.34	10,433.65
Surplus account, at end of year	$26,908.82	$16,965.34

Note: Includes $303.29 relating to receipts from 1952 meeting.

Schedule I

AMERICAN STATISTICAL ASSOCIATION

	Year ended December 31, 1953	1952
Publications:		
Journal—Printing	$15,118.77	$12,522.59
—Abstracts	750.00	
—Editorial expense	1,215.30	335.98
—Cost of old *Journals*	188.67	101.70
—Delivery charges	36.11	32.14
—Storage charges	114.00	96.00
	$17,422.85	$13,088.41
American Statistician	6,584.28	5,810.75
Acceptance Sampling	129.24	218.61
Membership Directory	831.86	1,300.00
	$24,968.23	$20,417.77
Miscellaneous Expense:		
Depreciation	$ 617.64	$ 555.87
Dues to other organizations	123.25	123.25
Bank charges	4.50	7.84
Taxes	314.89	559.41
Repairs and maintenance	153.25	206.95
Workmen's compensation insurance	24.53	26.91
Other	199.81	448.60
	$ 1,437.87	$ 1,928.83

SUMMARIES OF PAPERS DELIVERED AT THE 113th ANNUAL MEETING OF THE AMERICAN STATISTICAL ASSOCIATION IN WASHINGTON, D. C., DECEMBER 27 TO 30, 1953.

Edited by ARMEN A. ALCHIAN, *University of California (Los Angeles)*

The present section contains all available abstracts of papers presented at the 1953 national meeting of the American Statistical Association in Washington, D. C. The sequence of presentation here conforms to a grouping of abstracts according to the various sessions at which they were delivered.

PAPERS SUMMARIZED

ALLISON, HARRY, BRINSER, AYERS, AND ZWICK, CHARLES, *An Analysis of the Demand for Meat* . . . 362

ARROW, KENNETH J., *Development in Statistical Techniques of the Last Decade of Special Interest to Economists: Sequential Analysis* . . . 328

BACHRACH, CLIFFORD A., *Estimation of Length of Hospital Stay from Discharge Data* . . . 361

BANCROFT, T. A., *Organization and Functions of a Complete Centralized Statistical Center at a Land Grant College* . . . 332

BANCROFT, T. A., *Preliminary Tests and Pool Rules* . . . 348

BARANKIN, EDWARD W., *Theory of Behavior* . . . 364

BATES, GRACE E., *The Time-interval Approach to the Problem of Contagion* 368

BEALL, GEOFFREY, *Further Generalization of Neyman's Distributions* . . . 368

BECHHOFER, ROBERT, DUNNETT, CHARLES W., AND SOBEL, MILTON, *A Single-Sample, a Two-Sample and a Sequential Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances* . . 358

BELLOC, NEDRA B., *Validation of Morbidity Survey Data by Comparison with Medical Records* . . . 353

BERGER, MARTIN J., *Statistical Problems in Physics* . . . 369

BERKSON, JOSEPH, *Estimation of the Interval Rate in Actuarial Calculations; A Critique of the Person-Years Concept* . . . 363

BIRNBAUM, Z. W., *Characterization of Distribution-free Statistics* . . . 343

BOX, G. E. P., AND HUNTER, J. S., *The Study and Exploitation of Response Regions* . . . 328

BRETEY, PIERRE, *The Future of Railroad Shares* . . . 351

BRINSER, AYERS, ALLISON, HARRY, AND ZWICK, CHARLES, *An Analysis of the Demand for Meat* . . . 362

BROIDA, ARTHUR L., GEHMAN, CLAYTON, TRUEBLOOD, LORMAN, SCHWARTZ, M. H., MOSS, MILTON, CODY, PETER M., *Industrial Production Index* . . . 357

BROWN, ERNEST W., and WOODBURY, MAX A., *Time Series Factor Analysis with an Economic Application* . . . 331

BRY, GERHARD, *Trends and Cycles in German Wages* . . . 334

BURGESS, ROBERT W., *The Relation of Census Tracts to the General Census Program* . . . 367

BURR, IRVING W., *Use of Experiments in Engineering Statistics* . . . 358

CHERNOFF, HERMAN, AND LIEBERMAN, GERALD J., *A Note on the Use of Normal Probability Paper* . . . 330

CODY, PETER M., GEHMAN, CLAYTON, TRUEBLOOD, LORMAN, BROIDA, ARTHUR L., SCHWARTZ, M. H., AND MOSS, MILTON, *Industrial Production Index* . . . 357

COHEN, BERNARD M., AND COOPER, MAURICE Z., *A Follow-up Study of Mortality in World War II Prisoners of War* . . . 336

COLM, GERHARD, *The Economic Outlook for 1954* . . . 339

CONNOR, W. S., *New Experimental Designs for Paired Observations* . . . 360

COOPER, MAURICE Z., AND COHEN, BERNARD M., *A Follow-up Study of Mortality in World War II Prisoners of War* 336
COPELAND, MORRIS A., *Current Problems in Measuring Moneyflows* . . 363
CORNFIELD, JEROME, *Some Finite Sampling Concepts in Experimental Statistics* 369
CROCKETT, JEAN BRONFENBRENNER, *Effect of Current Operating Experience on the Realization of Investment Plans* 355
CROW, EDWIN L., *Statistical Determination of Tolerances in Rocket Development* 354
CUTTRELL, FLORENCE S., *Uses of Small Area Census Data in New York City* 365
CYERT, R. M., *The Use of Statistical Techniques in the Aging of Accounts Receivable* 359
DAILEY, JOHN T., *Achieving Maximum Prediction per Unit of Testing Time* 339
DALY, JOSEPH F., *Survey of the Theory of Finite Sampling* 369
DARROCH, J. G., *Role of a Centralized Statistical Organization in a University* 332
DAUER, ERNST A., *Limitations of Consumer Credit Statistics* 335
DAVIS, D. J., *Parametric Estimation of Survivorship* 361
DAY, B. B., *The Statistician in a Research and Development Laboratory* . 369
DEARDORFF, NEVA R., *Longitudinal Studies of HIP Experience, Background and some Findings of Pilot Study* 346
DECKER, RAY, AND FRAZIER, DAVID, *Gasoline Mileage in Winter Day-to-Day Use* 354
DEL PRIORE, F. R., AND KOMMERS, WILLIAMS J., *An Example of a Fractional Replication in a Bearing Abrasive Wear Test* 354
DOUTY, H. M., *Union Impact on Wage Structures* 334
DUGGAR, GEORGE, *Use of Census Tracts in Study of Changing Residential Patterns in Metropolitan Areas* 364
DUNNETT, CHARLES W., BECHHOFER, ROBERT, AND SOBEL, MILTON, *A Single-Sample, a Two-Sample and a Sequential Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances* 358
ECKLER, ROSS, AND TAEUBER, CONRAD, *The Current Statistics Program of the Census Bureau* 348
EISNER, ROBERT, *Expectations, Plans and Capital Expenditures: A Synthesis of Ex Post and Ex Ante Data* 356
ELY, J. EDWARD, *Adequacy of International Trade Statistics for Economics and Business Analysis* 350
ERHARDT, CARL L., *Hospital Morbidity Reporting—Experiences and Findings of a Pilot Project* 353
FOOTE, RICHARD J., AND KUZNETS, GEORGE M., *The Demand for Citrus Products* 361
FRAENKEL, MARTA, *Hospital Morbidity Reporting—Experiences and Findings of a Pilot Project* 353
FRAZIER, DAVID, AND DECKER, RAY, *Gasoline Mileage in Winter Day-to-Day Use* 354
FRYER, H. C., *The Role of a Centralized Statistical Organization in a College with Respect to the Reviewing of Manuscripts and the Promotion of Statistical Research* 332
GAFFEY, WILLIAM R., *The Problem of Within Family Contagion* 363
GEHMAN, CLAYTON, TRUEBLOOD, LORMAN, BROIDA, ARTHUR L., SCHWARTZ, M. H., MOSS, MILTON, AND CODY, PETER M., *Industrial Production Index* 357
GEORGE, JAMES P., *A Cyclical Index of Highway Ton Miles* 329
GHOSH, M. N., *Simultaneous Test of Linear Hypotheses by Analysis of Variance Methods* 348
GINZBERG, ELI, *Scientific and Professional Manpower* 327
GIVENS, MEREDITH B., *Tracts in Analysis of Worker Mobility* 366
GOLDFARB, NATHAN, *Longitudinal Study of Health Insurance Plan of Greater New York* 346
GOLDSTEIN, HAROLD, *Recent Advances in Statistics on Scientific and Professional Personnel* 327

GREENBERG, LEON, AND SEARLE, ALLAN, *The New Bureau of Labor Statistics Indexes of Productivity in Manufacturing* 331
GRIEVES, HOWARD C., *"Spot Checks" in Lieu of Complete Censuses* . . 347
GRIFFITHS, J. C., *Analysis of Variance Models in Sedimentary Petrology* . 344
GULBRANDSEN, R. A., AND MCKELVEY, V. E., *Problems in Sampling the Phosphoria Formation* 345
HANSEN, MORRIS H., AND HURWITZ, WILLIAM N., *Developments in Collection and Processing of Mass Statistical Data* 347
HARTLEY, HERMAN O., *Approximate Tests for Comparisons of Rank Correlations* 329
HELGER, BENGT, *International Criminal Statistics* 328
HOADLEY, WALTER E., JR., *Inadequacies of the Construction Estimates as General Economic Measures* 344
HODGES, J. L., *The Up-and-Down Method with Small Samples* 354
HORVITZ, DANIEL G., *On Respondent-Nonrespondent Differences Observed in the Pittsburgh Morbidity Surveys* 350
HOUSEHOLDER, A. S., AND KIMBALL, A. W., *A Stochastic Model for the Selection of Macronuclear Units in Paramecium Growth* 331
HUNTER, J. S., AND BOX, G. E. P., *The Study and Exploitation of Response Regions* 328
HUNTSBERGER, D. V., *An Extension of Preliminary Tests for Pooling Data* 348
HURWITZ, WILLIAM N., AND HANSEN, MORRIS H., *Developments in Collection and Processing of Mass Statistical Data* 347
HUTTON, THOMAS G., *Methodological Problems and Findings of Study of Recipients of Old-Age Assistance* 342
JONES, HOMER, *Recent Revisions of Consumer Credit Statistics* 335
JONES, HOWARD L., *Optimum Cluster Size* 329
KATZ, LEO, *Probability Distributions of Group Organization Theory* . . . 366
KIMBALL, A. W., AND HOUSEHOLDER, A. S., *A Stochastic Model for the Selection of Macronuclear Units in Paramecium Growth* 331
KOHN, ROBERT, *Questionnaire Design and Related Methodological Problems in the Canadian Sickness Survey* 351
KOMMERS, WILLIAM J., AND DEL PRIORE, F. R., *An Example of a Fractional Replication in a Bearing Abrasive Wear Test* 354
KRUSKAL, WILLIAM H., *The Problem of Non-normality, and Nonparametric Tests* 329
KUZNETS, GEORGE M., AND FOOTE, RICHARD J., *The Demand for Citrus Products* 361
LATHROP, JOHN B., *Operations Research as a Science* 355
LAURENT, ANDRÉ, *Use of Statistics in Engineering in France* 354
LEW, E. A., *Insurance Mortality Investigations of Physical Impairments* . 336
LIEBERMAN, GERALD J., AND CHERNOFF, HERMAN, *A Note on the Use of Normal Probability Paper* 330
LITTELL, ARTHUR S., AND SMITH, GEORGE V., *Intervals Between Onsets of Multiple Cases of Poliomyelitis in Families* 343
LOCKS, MITCHELL O., *Two Nonparametric Tests Using the Method of Ranks for Testing the Randomness of Samples Drawn from Finite Populations* 330
LOEWY, HARRIS, *The Flow of Net Cash Savings Through Life Insurance Companies* 362
LOOMER, HARLIN G., *Intercensal Needs for Small Area Data—By a Local Planning Agency* 357
LURIE, SIDNEY, *The Changing Stock Market* 352
MANDEL, JOHN, *Statistical Principles of Testing* 340
MARSHALL, HERBERT, *Problems of Co-ordination in the Canadian Statistical System* 342
MAUSER, FERDINAND F., *How the Automobile Industry Utilizes the Census Tract in Market Determination* 367
MEYER, HERBERT A., *Staffing a Central Statistical Organization* . . . 333
MICKEY, RAY, *On Testing the Association of Mineral Occurrence with a Set of Observable Characteristics* 345
MILLER, HERMAN, P., *Some Observations on the Inequality of Income* . . 337
MONTGOMERY, DOROTHY S., *Uses of Census Tracts by Housing People* . . 364

MOONAN, WILLIAM J., *Multivariate Analysis of Covariance for a Latin Square* 349
MOORE, FREDERICK T., *Some Statistical Evidence of Economies of Scale* 349
MOREHOUSE, M. DUTTON, *Trends in Monetary Structure in 1954* 352
MORGENSTERN, OSKAR, *Accuracy of Foreign Trade Statistics* 350
MORTON, J. E., *The Problem of Improving Mineral Statistics* 338
MOSES, L. E., *Some Comments on the Lot Plot Plan* 341
MOSHMAN, JACK, *A Two Sample Procedure for Linear Discrimination in Normal Samples* 330
MOSS, MILTON, GEHMAN, CLAYTON, TRUEBLOOD, LORMAN, BROIDA, ARTHUR L., M. H. SCHWARTZ, AND CODY, PETER M., *Industrial Production Index* 357
MUDGETT, B. D., *The Function of the Outside Consultative Committee in the Revision of Governmental Statistics* 336
MUNRO, S., *Some Aspects of Sequential Experimentation* 360
MYERS, PERRY H., *Use of Census Tracts for Business Analysis* 367
MYERS, ROBERT J., *Factors in Interpreting Mortality After Retirement* 337
MCKELVEY, V. E., AND GULBRANDSEN, R. A., *Problems in Sampling the Phosphoria Formation* 345
NETER, JOHN, *Problems in Experimenting with the Application of Statistical Techniques in Auditing* 358
NUTTER, G. WARREN, *Monopoly, Bigness, and Progress* 352
OLDS, EDWIN G., *Using the Experimental Approach in the Teaching of Statistics* 358
PHILLIPS, A., *The Stability of Technical Coefficients: Evidence from Inter-Plant Differences in Labor and Materials Productivity* 350
PRESTON, G. W., *Contemporary Topics in Statistical Physics* 368
QUACKENBUSH, G. G., *Demand Analysis from the M.S.C. Consumer Panel* 362
REID, A. T., *Stochastic Processes and the Study of Growth Phenomena* 331
REID, MARGARET, *Value of Dwellings in Relation to Income* 337
RICE, STUART A., *Problems of Coordinating the United States Statistical System* 341
ROMIG, HARRY G., *Probability of Acceptance for Sampling Plans Based on Average-Standard Deviation Acceptance Criterion* 330
SCHRADER, WILLIAM B., *The Validation of Testing Programs for University Students* 340
SCHWARTZ, M. H., GEHMAN, CLAYTON, TRUEBLOOD, LORMAN, BROIDA, ARTHUR L., MOSS, MILTON, AND CODY, PETER M., *Industrial Production Index* 357
SEARLE, ALLAN, AND GREENBERG, LEON, *The New Bureau of Labor Statistics Indexes of Productivity in Manufacturing* 331
SELLIN, THORSTEN, *Problems and Prospects of Criminal Statistics in the United States* 331
SHRYOCK, HENRY S., JR., *Changing Geographic Patterns of Migration in the United States* 338
SIMMONS, W. R., *Business Uses of Intercensal Data* 357
SMITH, GEORGE V., AND LITTELL, ARTHUR S., *Intervals Between Onsets of Multiple Cases of Poliomyelities in Families* 343
SMITH, HARRY, JR., *Weighting Coefficients for Age-Adjusted Death Rates* 349
SOBEL, MILTON, BECHHOFER, ROBERT, AND DUNNETT, CHARLES W., *A Single-Sample, a Two-Sample and a Sequential Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances* 358
STRODTBECK, FRED L., *Family Interaction and the Transmission of Achievement-related Attitudes* 340
TAEUBER, CONRAD, *Meeting the Needs for Small Area Intercensal Data* 357
TAEUBER, CONRAD, AND ECKLER, A. ROSS, *The Current Statistics Program of the Census Bureau* 348
TAYLOR, WILLIAM F., *Consistency of Estimators under a Specialized Bioassay Procedure* 364
TRUAX, DONALD, *An Optimum Slippage Test for the Variance of K Normal Populations* 359

TRUEBLOOD, LORMAN, GEHMAN, CLAYTON, BROIDA, ARTHUR L., SCHWARTZ, M. H., MOSS, MILTON, AND CODY, PETER M., *Industrial Production Index* 357
TUKEY, JOHN W., *Unsolved Problems of Experimental Statistics* 343
WAGENHALS, R. E., *Demonstration Teaching of Statistics* 358
WALLIS, W. ALLEN, *Responsibility of a Centralized Statistical Organization in a University for Training in Statistics* 333
WALLIS, W. ALLEN, *Teaching Statistics to Executives* 356
WEINGARTEN, HARRY, *Continuous Sampling Plans* 341
WELLS, ORIS V., *Economic Forecast of the Agricultural Situation, 1954* 339
WENTWORTH, EDNA C., *Methodological Problems and Findings of Survey of Aged Beneficiaries of Old-Age and Survivors Insurance* 342
WOLMAN, LEO, *Wages Since 1914* 334
WOODBURY, MAX A., *Information Theory and Prediction* 360
WOODBURY, MAX A., AND BROWN, ERNEST W., *Times Series Factor Analysis with an Economic Application* 331
ZWICK, CHARLES, BRINSER, AYERS, AND ALLISON, HARRY, *An Analysis of the Demand for Meat* 362

Scientific and Professional Manpower: the Value and Limitations of a Statistical Approach. ELI GINZBERG, *Columbia University.*

A major shortcoming of American social science is its infatuation with the doctrine that if one only had a sufficient number of facts, one could solve any problem. This particular preconception is grounded in the following: (1) Our pragmatic position with its anti-theoretical bias. (2) The availability of funds to support expensive data collection undertakings. (3) Our desire to make progress quickly and our naive belief that the activities connected with data collecting are proof of progress. (4) Our belief that the "facts" will be able to resolve all difficulties and that there are no underlying value conflicts.

During the past few years much energy has been poorly directed because of an error in strategy: We started to collect large bodies of data without knowing what to do with them, and without having any clear idea of the key questions they were supposed to answer. Much of this data collection has been spearheaded by interested pressure groups to support a particular point of view. Much of the governmental and academic effort in the collection of data has been of dubious value. One of the most serious results of a preoccupation with statistical tabulations has been the blithe assumption that the concept of shortage (or balance) in professional manpower is a simple arithmetic relation between supply and demand. Similarly, there has been no proper attention paid to the role of substitution in matters of supply. Among the other significant facets of the problem that are not illuminated easily via statistics is that of utilization, which can of course greatly influence whether any given supply proves to be adequate or not. Entirely too little emphasis has been placed upon the fact that any particular manpower problem can probably be resolved in any one of a series of ways depending in large part upon the criteria that are employed. Perhaps the most serious shortcoming of all growing out of a quantitative approach is the inherent tendency contained therein to gloss over qualitative differences among professional persons and to see the problem primarily as one of numbers.

There are at least five major areas in which organized statistical efforts can make a significant contribution to the illumination of scientific and professional manpower problems: (1) By providing knowledge of the occupational structure; (2) By facilitating studies of the probable size of future supply; (3) Through contributing to the organized study of future demand by studying the strength of particular factors that have influenced demand in the previous time periods; (4) By helping to set out in a systematic fashion the incentive factors in different occupations which influence training and distribution of trained persons; and (5) by studies of the flow of individuals in and out of different types of employment, the strategically important question of convertibility of trained manpower can be illuminated. The rate at which a statistical approach can contribute along the foregoing lines to the understanding and solution of professional manpower problems will depend very greatly on the extent to which existing theory can be improved with respect to the concept of balance, the potentialities and limitations of "convertibility" in the use of highly trained persons, the elaboration of the significant relations between qualitative and quantitative considerations, and the factors determining different utilization levels.

Recent Advances in Statistics on Scientific and Professional Personnel. HAROLD GOLDSTEIN, *Bureau of Labor Statistics.*

Widespread interest in professional personnel has focused attention on statistical data in this field. Recent improvements include (a) improvement in the frequency, currency, detail, and over-all quality

of the data on output of trained personnel from the educational institutions; (b) development of techniques for registering members of professions; (c) active participation by professional societies in studies of their fields; (d) a beginning in the study of occupational mobility in the professions. The kinds of data we have and the gaps in information reflect the emphasis on large-scale statistical surveys; more intensive studies are needed to provide missing data.

Among the areas of work which should be productive are: I. Development of measures of qualitative differences among individuals to temper the purely quantitative information now available. II. Development of surveys of professionals via employing establishments, as an essential supplement to census-type and professional society membership type surveys. III. Investigation of factors and concepts underlying demand for professional personnel. The study of professional personnel within the context of general personnel and manpower problems is essential.

International Criminal Statistics. BENGT HELGER.

In spite of repeated efforts over a hundred years to achieve greater uniformity of statistics on crime, a direct comparison of national series showing crime rates, and the trend of these rates, is scarcely possible. Criminal statistics cannot be taken at their face value, and international comparisons are particularly falacious without sufficient knowledge of all particulars, pertaining to the penal and judiciary system of countries, that are essential for an interpretation of national figures on crime. A new approach to the problem of international comparisons would consist in selecting for this purpose certain actions that are universally recognized as criminal and are of such a nature that they regularly become known to the authorities, and redefining these actions so as to give a uniform content to the series that are compared. Such data, stemming from the police or the investigating authorities, would be valuable as indicators of the frequency of crime, whereas court statistics, supplemented by information on suspended sentences, probation and parole, etc., have their value for studies of the treatment of offenders. Current statistics on crime can scarcely reveal the causes of criminality but may serve as a frame for further research into this subject by giving a general picture of the dimensions of criminality, the incidence of various types of criminality within socio-economic groups, and its evalution under the impact of social changes. In this connection, international comparisons present a great deal of interest.

The Study and Exploitation of Response Regions. G. E. P. BOX and J. S. HUNTER.

Techniques have been proposed by Box and Wilson for the study of response surfaces. Having reached a near-stationary region in the factor space, it is of extreme practical importance to study the local features of the response surface and to exploit the surface to best advantage. Designs are needed which will allow the fitting of a polynomial of degree d as efficiently and economically as possible. Optimal designs of first order (for the fitting of plane surfaces) have been described by Box elsewhere. The present paper discusses the problem of deriving optimal designs of second and higher order. It is shown that the principle of obtaining designs giving smallest variances for the estimates does not in the case of second and higher order designs result in a unique or a necessarily desirable solution of the problem. Some other principle seems to be required. The concept of information distributions is introduced. This distribution yields the information per observation obtained at a point on the fitted surface at any location in the factor space.

It is suggested that a *spherical* information distribution is desirable (i.e., a distribution such that the information is constant on spheres centered at the origin of the design). Designs which give a spherical information distribution satisfy the criteria that the variance-covariance matrix and the moment matrix of the design are invariant when the design is rotated. These new designs are therefore called "rotatable" designs. It is shown that the principle of rotability leads to optimal designs of first order. Second order rotatable designs are then derived. The designs are also considered from the point of view of the biases in the estimates which may be present if the assumed model is inadequate. Some confounding systems are discussed.

Developments in Statistical Techniques of the Last Decade of Special Interest to Economists: Sequential Analysis. KENNETH J. ARROW, *Stanford University.*

This paper is a survey of the fundamental ideas of sequential analysis and its implications for economics. The general conclusion is that sequential analysis as a statistical technique is likely to find few applications to econometrics because of the special problems of sampling in the typical problems of economic statistics, but that the type of reasoning underlying sequential analysis is likely to be of considerable importance in the theory of dynamic economics, particularly where actions with random consequences are under consideration. The technique of sequential analysis of statistical data is applicable when the experimenter can control the taking of additional observations on the basis of the results of earlier ones. The additional observations are possible at a cost. The technique is therefore not applicable to time series already collected, data collected as a by-product of other activities, to surveys

where it may be difficult to enumerate after each observation, or to multi-purpose surveys, where the stopping-point may be different for different questions asked. A brief development of the theory of sequential decision between two hypotheses is given from a Bayesian viewpoint. After a number of observations, the original *a priori* probabilities are transformed into a set of *a posteriori* probabilities, but the situation is essentially unchanged, except for the magnitudes of the probabilities. This simple observation enables an immediate derivation of the sequential probability-ratio test. A problem of choice among inventory policies is then stated. It is shown that the same type of reasoning used in the analysis of the sequential decision problem leads to a functional equation which can be solved to yield an optimal solution to the inventory problem.

The Problem of Non-normality, and Nonparametric Tests. WILLIAM H. KRUSKAL, *University of Chicago*.

Most statistical procedures used in practice are based upon the assumption of normality. The arguments usually given in defense of this assumption are stated and critically discussed. A method of circumventing the assumption of normality is the use of nonparametric statistical procedures. The advantages and disadvantages of these procedures are discussed in general, and some examples are given. The procedures of applied nonparametric analysis are classified, and key references are given.

A Cyclical Index of Highway Ton Miles. JAMES P. GEORGE.

This investigation has as its objective the construction of a composite weighted cyclical index of ton miles operated by trucks and truck combinations on all main and local rural roads in the United States. The index computes for any particular month the cyclical position of highway ton miles with respect to the computed normal or expected value as of that month. Polynomial curves of appropriate degree were fitted to the cyclical-irregular movement of highway ton miles over a 78-month period. Criteria, affording a test for goodness of fit, include the variance ratio or Snedecor's "*F*," the standard deviation of the residual variance, and the appearance of the plotted curve. This fitting (or smoothings process eliminates the irregular movements and causes those fluctuations attributable to cyclical force) to stand out in bold relief. It should be pointed out that the cyclical irregular movement to which the curve is fitted is in terms of composite weighted standard deviation units from "normal."

In consequence of this investigation the writer has been able to isolate and to delineate a structural pattern of the cyclical movement of highway ton miles which, with varying intensity, tends to repeat itself at approximately 32-month intervals. This mathematical function (the orthogonal polynomial curve) appears to postulate an underlying law, governing the behavior of highway ton miles. The structural pattern of the cyclical movement of highway ton miles is predicated upon an assumption of economic rhythm in the ton mile series. Even though trend and/or seasonal do change, the structural pattern of highway ton miles can be expected to continue, showing much the same contour as before. In other words, the contour of the structural pattern will tend to remain more or less constant irrespective of the level from which measured. From the standpoint of forecasting this is a most important consideration.

This investigation has been in part analytic and in part synthetic. In other words, it first concerned itself with a breaking down of the highway ton mile series into its component elements. In the latter stage, the investigation directed its efforts to a recombining of the constituent elements into a theoretical ton mile series. The synthesis makes possible an extrapolation or projection into the future. It also serves to verify or confirm the correctness of the analysis by reconstructing the highway ton mile series from its constituents.

Approximate Tests for Comparisons of Rank Correlations. HERMAN O. HARTLEY.

Of the numerous measures of rank correlation the two best known are: Spearman's rank correlation r_s and Kendall's rank correlation t_k. Some results on the exact distribution of these measures are known for the case of 0-correlation and some moments for the case of ranks generated by a bivariate normal population with correlation coefficient ρ. In the latter case the present note investigates the z-transformation or r_s and t_k, both by the method of statistical differentials and by Monte Carlo calculation. It is found that for moderately large sample size both z-tranforms are approximately normal with variances approximately independent of ρ. This property results in simple tests of significance for comparisons of rank correlations so transformed. The restriction to ranks generated by samples from bivariate normal samples is lifted and the tests shown to apply to a much wider class of ranks.

Optimum Cluster Size. HOWARD L. JONES.

When a two-stage sampling procedure consists in selecting a number of clusters of equal size, and the total cost is a linear function of the number of clusters and the number of individuals selected, the optimum cluster size is a simple function of the parameters of the cost function and the intraclass and interclass correlation coefficients. Suggestions for estimating these parameters and coefficients are

proposed and discussed for an illustrative example. The more general situation is then examined where the number of attempted selections is the same for every cluster, but the actual number of individuals inspected varies from cluster to cluster.

A Note on the Use of Normal Probability Paper. HERMAN CHERNOFF and GERALD J. LIEBERMAN.

This paper illustrates, with a special example, that the graphical technique to be applied to a problem should depend to a large extent on the use to which the graph is to be put. In particular, we treat the problem of selecting the representation of a sample on normal probability paper when it is desired to obtain "optimum graphical" estimates of the mean ξ and standard deviation σ of a normal distribution.

Two Nonparametric Tests Using the Method of Ranks for Testing the Randomness of Samples Drawn from Finite Populations. MITCHELL O. LOCKS, *University of Oklahoma*.

In this paper are developed two nonparametric tests based on the method of ranks, both of which may be used to test the randomness (or representativeness) of samples drawn from finite populations. The theory behind the development of these tests is essentially the same as that governing tests for randomness which use the original values. Since the application of these tests requires the ranking of all of the items in the population according to some numerical characteristic, the tests can be applied only if values with respect to at least one numerical characteristic are available for every single item in the population. Admittedly, this restricts the scope of usefulness of these tests. However, it is believed that the tests can be used more profitably for certain types of problems (e.g. tests for randomness of samples drawn from J-shaped and U-shaped populations) than other statistical tests for randomness now in use.

A Two Sample Procedure for Linear Discrimination in Normal Samples. JACK MOSHMAN, *Oak Ridge National Laboratory*.

The problem of discrimination from a normal linear regression model considered is that of obtaining a confidence interval for x corresponding to an observed y when the expectation of y, given x is $E(y|x) = \alpha^* + \beta^* x$. Various incongruities are discussed which result from a naive approach. A deterministic one sample procedure is shown to have two difficulties: (1) The procedure is inefficient in that the confidence interval obtained is larger than a specified $1-\gamma$; (2) The excess is a function of the unknown variance. A two sample procedure is exhibited which makes the excess over $1-\gamma$ to be less than a specified $\delta (0 < \delta < \gamma)$ whenever $|\beta^*| < l$ for any predetermined $l > 0$ and the excess is independent of the variance.

Probability of Acceptance for Sampling Plans Based on Average-Standard Deviation Acceptance Criterion. HARRY G. ROMIG.

The evaluation of a variables sampling plan using Average-Standard Deviation Acceptance Criterion requires Probability of Acceptance values for determining its Operating Characteristic. The mathematical relations are presented for computing such probabilities as developed in 1934 assuming a normal law distribution for the parent population and resultant exact distributions for averages and standard deviations where the correlation between averages and standard deviations is zero for two cases: 1. Sample size n large: distributions of averages and standard deviations assumed normal; and 2. Sample size n small: distribution of averages normal and of standard deviations non-normal. These results may be compared with later work started in 1947 at Stanford University by Goode, Bowker, Ireson, and Resnikoff.

For the two cases a cutting plane slices the frequency surface for averages and standard deviations and volumes under the surface are determined for shifts in level (average), variability (standard deviation), or both, due to changes in the system of causes. Levels of incoming quality with respect to engineering requirements and the protection desired for any sampling plan determine the location of the cutting planes used for evaluation. The Average-Standard Deviation Acceptance Criterion is set up for either a minimum or maximum engineering limit for sample values of averages and standard deviations. For such criteria, relations are provided for determining the volume under the surface to the right or left of the cutting plane covering both max. and min. limits. These relative volumes provide the desired probability of acceptance values for any postulated incoming quality p'. A simple approximation for n small is given. Exact and approximate P values for samples of 5 are presented for two sampling plans for evaluating the error of the approximation. A graphical solution of the multiple integrals is described. These probability of acceptance values must be obtained in order properly to set up various types of variables sampling plans.

Time Series Factor Analysis with an Economic Application. ERNEST W. BROWN and MAX A. WOODBURY

It has been recognized that the application of many statistical techniques to time series is invalid due to the correlation of successive observations. Among the techniques affected is factor analysis, since many of the underlying relations, which a factor analysis should uncover in a time series, act with a time delay, and hence would not be discovered. An appropriate definition of a factor and a technique of time series factor analysis is formulated and discussed and an application to a set of ten economic time series is made. The ten series cover national product, income and employment. In factor analysis, if there are n original variables, no more than about $n/2$ factors can be identified. An interesting feature of time series factor analysis is that this does not necessarily hold true. In fact, factor analysis of time series may require, and allow identification of, more factors than variables.

A Stochastic Model for the Selection of Macronuclear Units in Paramecium Growth. A. W. KIMBALL and A. S. HOUSEHOLDER, *Oak Ridge National Laboratory.*

Prior to division, the macronuclear units in *Paramecium* (see, for example, Sonneborn, *Ann. Rev Microb.*, pp. 55–80, 1949) presumably double in number in a manner similar to chromosome doubling at mitosis. During the process of division each daughter animal receives approximately one-half of the units present. Thus in any population of *Paramecia*, each animal has about the same number of units. If there are n units per animal, there may be as many as n different types of units, or all units may be alike, and any combination between these extremes is also possible. Such combinations are called states. Given the state of a single animal and given an hypothesis about the selection of macronuclear units by daughter cells, the probability distribution of states after N divisions may be computed readily by the methods of stochastic processes. Results have been obtained for the hypothesis of completely random selection and have been compared with experimental data.

Stochastic Processes and the Study of Growth Phenomena. A. T. REID, *Columbia University.*

This paper is divided into three parts: (1) Introduction, (2) Construction of stochastic models, and (3) Statistical inference in stochastic models. In Part 1 we discuss the deterministic and stochastic approaches to the study of growth phenomena; and give an introduction to the theory of branching stochastic processes as developed by Bellman and Harris. We consider in Part 2 the construction of various stochastic models for growth using the above theory. Models for birth, birth-and-death, and mutation processes are discussed. The use of these formal models in the study of epidemics and rumor spread, as well as in the study of bacterial growth, is pointed out. In Part 3 we discuss problems of estimation and testing associated with stochastic growth processes. Previous work is reviewed, and some recent investigations on sequential decision problems for branching processes is discussed.

The New Bureau of Labor Statistics Indexes of Productivity in Manufacturing. LEON GREENBERG and ALLAN SEARLE

The Bureau of Labor Statistics is planning to publish several series of productivity indexes in manufacturing covering the years 1939, and 1947 through 1952. Two of the series will show changes in man-hour requirements for physical production in manufacturing. A third series, based on concepts similar to the gross national product approach, will show changes in man-hour requirements for dollar value added in manufacturing and will reflect the influence of changing inputs of material, supplies and fuel. The various measures fill in a statistical gap of 13 years for which reliable all-manufacturing productivity statistics are not available.

The indexes are being derived from secondary sources of production and man-hour data. In relating these two statistics to derive productivity ratios numerous data problems are encountered. Not all of them can be solved immediately and the indexes released by the Bureau will not be precision instruments. But they will still be useful and necessary indicators in the broad areas of economic and business analysis, manpower planning and productive efficiency.

Problems and Prospects of Criminal Statistics in the United States. THORSTEN SELLIN.

Criminal statistics issued by agencies dealing with crime or criminals have been assumed to be important sources of data for the scientific study of the trends of criminality, the characteristics of offenders and the efficiency and effectiveness of the agencies dealing with them. The dearth of good criminal statistics in the United States is discussed and reasons for this situation considered. The main part of the paper examines the problem of how such statistics can be used in the study of trends of criminality and the characteristics of the offender, and raises a number of questions designed to stimulate a discussion of this problem. Illustrations are offered pointing to the need of improving certain official reports now issued, attention being drawn in particular to the need for some revision of the standard classification of offenses employed by many national and state agencies and the need for more carefully computed crime rates than those now being published.

Organization and Functions of a Complete Centralized Statistical Center at a Land Grant College. T. A. BANCROFT, *Iowa State College.*

The position is taken that there is no unique optimum organization or program for a statistical center which will work equally well at all universities or even at all land grant colleges. Further, even at the same institution either one of several alternative arrangements might work equally well. A *complete* statistical center should provide: (i) a research and teaching program in statistics per se in order to develop new statistical theory and methodology and train statisticians; (ii) a service teaching program to provide for basic general courses in theory and methods and specialized courses in statistics for students majoring at the undergraduate and graduate levels in some other substantive subject matter area; (iii) a consulting service program, i.e. recognized and budgeted time for various staff members of the statistical center to consult with research workers on investigations involving the use of statistical theory and methods; and (iv) a computing service for the programming and analyses of data resulting from research investigations. These four main objectives could be accomplished by a separate department of statistics, or, in case of a small group, a sub-department, and a campus-wide statistical laboratory of institute status.

Role of a Centralized Statistical Organization in a University. J. G. DARROCH, *Washington State College.*

The recommended relationship between the statistical organization and the various research units of the University may be set forth as it affects some of the major units. (1) The Agricultural Experiment Station—two or more staff members assigned to service this unit, with the privilege of calling on their associates to aid in special problems or to share the load at seasonal peaks. The consultor should be permitted to exhibit a preference among the possible consultants, if he so wishes. Some Directors might prefer to have the consultants attached to their staff but the preference would seem to be to house them with the statistical group, unless there is a distance problem, and have them administratively responsible to the head of this same group. (2) The Department of Mathematics—the liaison here, established by joint appointment or some other effective means, needs to be rather close. The relationship here on consulting problems may well be one of the statistician consulting with the mathematician. (3) The Social Science Departments—the needs here could best be served by staff members specifically trained in certain statistical disciplines, again a matter of assignment. (4) Engineering, Medicine and other fields—while specialized knowledge would be an advantage there are many problems where a straightforward statistical approach would be of material assistance to the research program. The inference here is that much can be done by the general consulting staff in the early development of a consulting load. As the loads from these areas build up it would be advisable to add personnel with background training in the various fields. (5) The computing facility—for statistical problems the consultant should provide the liaison between the worker and this unit. Thus it would appear desirable that the computing unit be an integral part of the organization under discussion.

The other aspect of the consultation program that should be mentioned is that having to do with the needs of graduate students. The most desirable pattern would seem to be that of the graduate student and his major professor both participating at least in the initial consultation. At that time the relationship of the consultant to the research problem should be fairly clear, and responsibility for direction of statistical matters and for computing decided upon.

The Role of a Centralized Statistical Organization in a College with Respect to the Reviewing of Manuscripts and the Promotion of Statistical Research. H. C. FRYER.

Statistical methods should not be forced on the research staff of the college or university, but the arguments for their use should be brought to the attention of the administrators of the research programs. To that end the following system is recommended: (a) Every proposed new project should be studied by a committee including a representative of each department involved, the statistical group being considered to be involved whenever the proposed experiment is to produce sampling data. (b) This committee is to make written recommendations to the directors of the research programs, including the presentation of minority recommendations when decisions are not unanimous. (c) When manuscripts result from the research they are subject to the rules of (a) and (b).

Most colleges and universities have limited funds, and someone must be convinced that the research will produce the greatest good for the money expended. The consumable end-product is some form of applied statistics; hence, it is suggested that every statistician in the organization should divide his or her time among teaching, statistical consultation with those doing applied research, and pure statistical research. The division should not necessarily be in thirds; emphasis should be on the activity best adapted to each person's interests and abilities. This plan should point statistical research toward problems the applied scientists need to have solved, and should keep the statisticians from getting into intellectual "ruts." Some may think the cost of paying several persons capable of research instead of one or two who do nothing else will be greater. Many of our young Ph.D.'s are capable of research,

and would gain valuable experience in such positions. If they like this multi-purpose activity and choose to stay with it, their productivity should increase with the cost of keeping them indefinitely. Before they settle permanently on such a job they will realize there is a limit to the pay they can receive; and they can go elsewhere if they prefer. In the meantime the school will benefit by the newness and the vigor of their statistical activities. Those schools in which money is no limitation and which do not have to worry about the applications of their research products may wish to supplement the above plan with some eminent research personnel who wish to do full-time theoretical research.

Staffing a Central Statistical Organization. HERBERT A. MEYER.

The number and qualifications of the members of the staff of the Central Statistical Organization depend upon the functions and responsibilities that are assigned to the Central Organization. These might be grouped under the following heads: (1) Consultation, (2) Teaching—theory and application of statistical methods, (3) Evaluation of manuscripts where statistical methods are used, (4) Research, theoretical and applied, and (5) Computing Service (machine room).

In a large university, the technical staff of the Central Statistical Organization should be composed of specialists in various statistical fields. Where teaching is one of the functions, the staff will have to be quite large and those who teach should also do research and consultation. Where teaching is decentralized, joint appointments of at least some of the teachers would seem to be desirable in order to bring the Central Organization in close relationship with those departments where problems in the application of statistical methods generally arise. The plan of joint appointments is working quite satisfactorily in the University of Florida, where the teaching of statistics is not centralized.

In a small university, the same general organization would seem desirable, but if funds are limited and demands less varied, the staff will be small. This calls for greater versatility on the part of the members of the staff While in a large university the Central Statistical Organization may be, and generally is, an independent unit, in a smaller institution it may be necessary for financial or other reasons to make it a branch of the mathematics or some other department. The computation service, or machine room, is an important section of the Central Statistical Organization. To be effective, it must be staffed with a competent supervisor familiar with statistical methods, as well as with competent keypunch operators and machine operators. Where courses are given in the university in the operation of IBM and other types of computing machines, these can serve to train potential replacements and additions to the card machine computing section. It is a difficult problem to find persons with the desired qualifications who are willing to accept the starting university salaries because of the demand in industry and government for statisticians at salaries above those in most universities; and this, the problem of finding a staff at all, is perhaps the most difficult of all staffing problems.

Responsibility of a Centralized Statistical Organization in a University for Training in Statistics. W. ALLEN WALLIS.

Departments in fields other than statistics normally do not have faculties qualified to teach current statistical methods, so the responsibility for statistics courses should not rest with them. On the other hand, statistics departments vested with sole authority over the statistics courses may tend to overemphasize statistics and its mathematical prerequisites at the expense of the students' subject matter studies. At Chicago, we have developed what looks to be a promising method for reducing duplication of statistics courses in the various departments and raising the level of instruction, while avoiding the pitfalls of authority centered in the statistics department.

Our first step was to introduce a new elementary course taught by one of our own men. After this course was well developed, people teaching statistics in various departments were invited to merge their course with ours. In general, their reaction was to introduce our syllabus and other materials but retain their instructors and the independence of their courses. The syllabus and teaching materials were revised somewhat in the light of suggestions and criticisms from these instructors in other departments, and then one or two departments did merge their courses with ours. As those departments reported favorably on their experience, others took the plunge. Essentially the same evolution has occurred with the advanced course on statistical inference.

With this system, each department lists the unified course in its own section of the catalog with its own number; but all designate the same room and instructor. Any department could secede simply by making a different room assignment in its Announcement. Actually, this would not be likely to happen without consultation with the statistics department, since other departments know that we are sensitive to their needs and limitations as well as to our own standards and requirements. Unification of statistics courses prevents their being too closely tied to a narrow range of substantive problems, encourages the use of examples from a variety of fields, and is generally an advantage pedagogically.

Whether the device used at Chicago would work at other universities is a question that must be answered in the context of each institution. If, with the elimination of a course to reduce duplication,

a department suffers cuts in appointments and budget, then it will be less ready to reduce the duplication. Some budgetary readjustments may be in order; but so long as cuts are not allowed to work to the direct disadvantage of those who help achieve the gains, many of the obstacles to eliminating duplication in universities seem to be avoided.

Trends and Cycles in German Wages. GERHARD BRY.

The paper compares German, English, and American long-term trends in wage levels and in wage structure, as well as their cyclical behavior, from 1871 to 1953. Hourly money wages are shown to have increased about seven fold in Great Britain, eightfold in Germany, and about ten fold in the United States. The divergence in the increase of real wages—two and one-half fold for Germany and Great Britain and fourfold for the United States—indicates the close relationship between wages and economic and political fortunes in these countries. A tendency toward decreasing differentials can be observed in the three countries. This trend towards greater equality is relatively clear in German skill, sex, age, and regional differentials. It is suggested also in city size and industrial differentials. Basically, this trend follows from the industrialization process itself, which leads to reduced differences between wage earners.

Although wage rates respond to general business contractions—by decreasing rates of growth, stagnation, or declines—there is disclosed a definite downward rigidity of wage rates. During the period 1871–1953 only two substantial declines in German money wage rates occurred. Earnings have larger cyclical amplitudes than rates because overtime, output bonuses, and hours worked are cyclically more sensitive. Even in the Great Depression, real weekly earnings in Germany and in the United States declined by only 15 per cent. In both countries the real sufferers from the Great Depression were the unemployed rather than the employed workers.

Union Impact on Wage Structures. H. M. DOUTY, *Bureau of Labor Statistics.*

A series of studies by the Bureau of Labor Statistics indicate that relative wage differentials among occupations have declined markedly in American industry, particularly during recent years. Over the past half century, several factors calculated to narrow job differentials in the long run can be distinguished. Among these are the sharp decline in immigration during World War I and the subsequent adoption of a restrictive immigration policy; a declining birth rate until the 1940's; a rising level of education and training among the working population; and the mechanization of large areas of unskilled work. These broad labor market forces clearly produced some tendency for relative job differentials to diminish. The decline in job differentials during the past decade, however, was much sharper than would have been anticipated on the basis of long-run labor market forces alone. One factor was governmental wage policy, notably in the economic stabilization program during World War II. Of major importance was a marked tendency for unions, beginning in the defense period in 1941, to formulate their wage demands in terms of uniform money increases (and hence unequal percentage increases among jobs) and settlements to be made in this fashion.

Among the reasons for this development were: (1) uniform money increases tend to appear more equitable in an inflationary period when wage increases are designed largely to offset increases in living costs; (2) it is politically easier, especially in an inflationary situation, for union leadership to press for uniform money increases; (3) within limits, skilled workers may be content with the maintenance of absolute wage differentials, even though their relative wage position is deteriorating; (4) on the whole, employers during the past decade have seemed more concerned with the size of negotiated wage increases than with the form of their distribution; (5) the structure of wages probably seems less important than the general level of wages, at least in inflationary periods. The decline in relative wage differentials among jobs has been much more marked in some industries than in others. Divergence in experience within manufacturing is well-illustrated in the basic steel and automobile industries. Assuming reasonable economic stability, it is probable that relative job differentials will receive considerable attention during the next few years from unions and employers. Special wage adjustments for skilled workers in some industries during 1953 point in this direction. Systematic review of the structure of job rates may occur in a significant number of situations.

Wages Since 1914. LEO WOLMAN, *Columbia University.*

This paper deals with one segment of wages in this country, the behavior of hourly earnings in five industries—manufacturing, class I railroads, building, anthracite and bituminous coal mining. Average hourly earnings, gross and net, are presented as an approximation of changes in the price of labor during these 40 years. Despite interruptions in their upward movement, money and real hourly wages multiplied many times over the period. All five groups shared in the rise, but unequally. Anthracite coal topped the list with nearly a 10-fold increase; building trades' rates rose least, less than 6-fold. Real hourly wages likewise mounted, but, of course, less than money wages. Anthracite real wages rose

3.5 times, manufacturing 2.8, and building over 2 times. The interesting question is, what economic conditions accounted for this multiplication in real and money wages? Using manufacturing wages as an illustration, it is clear that the bulk of the total money advance of $1.47 an hour was a product of the economic conditions of war and postwar boom, since only 10 cents of this amount were added between 1920 and 1940.

When money wages are converted into real wages, surprising results are obtained. For real wages showed marked improvement without apparent regard to movements in money wages, business conditions or union organization. Comparing the course of money and real wages in the two wars one encounters many difficulties. The first was shorter and didn't have wage controls. In several industries wages in World War II were fixed in long-term contracts. Bearing these factors in mind, the available evidence suggests that money and real wages rose more between 1916 and 1918 than from 1941 to 1945. This record of 40 years suggests that the period beginning in 1914 can be described as a high-wage era in the sense that the increases of this period were not matched by the wage advances of the preceding half-century. On the question of organized labor's effect on wages, the evidence is conflicting. Attention should be called, however, to the striking rise in manufacturing real wages between 1929 and 1940 when average unemployment stood at a consistently high level and when the only tenable explanation of the rise in real wages must be public policy and trade unionism.

Limitations of Consumer Credit Statistics. ERNST A. DAUER, *Household Finance Corporation.*

Definition and content. All elements in the Federal Reserve Board's definition are relatively simple and appear clear. However, differences arise in attempting to apply the definition, because there are legitimate differences in point of view, and because different people expect the figures to serve different purposes. Even the Federal Reserve Board does not seem to appreciate fully that difficulty also arises because the consumer credit total figure is used, as a rule, for an altogether different type of analysis than the break-downs. Their decisions on the items to be included for one type of analysis make it impossible for the resultant figures to serve properly in another analysis. The basic published estimates are limited to the amounts of credit outstanding. Figures showing the flow of consumer credit, if added, would provide a much more complete understanding of current developments.

Methods of estimation. The recent revision has distinctly improved the technical accuracy and coverage of the estimates. Even with their shortcomings, they are substantially better than the statistics available in most economic areas. I am concerned about the broader, non-technical aspects of the problem. Trained statisticians are aware: that *all economic* statistics measure differences of degree; that all classification is subjective; and that every problem really requires its own set of statistics, if they are to be truly germane. Most users, lacking formal training, are misled by the aura of accuracy and precision which surrounds the mere issuance of estimates by a federal government agency. A clear, concise, and simple disclaimer is necessary to dispel the mirage of infallibility.

On the surface the estimates provided for commercial banks—the most important single type of holder—appear to be the most thoroughly grounded of any of the series. Yet, the whole procedure involves guesswork, because the banker frequently does not know the use of each loan, at the time it is made; and experience shows that many respondents give whatever figures are easiest when condition reports are prepared, semi-annually. The correction factor applied by the Board to exclude the non-consumer portion from the reported figures is based upon a single survey—a sample which seems wholly inadequate. The validity of using the findings of this single survey from 1939 to the present time, and into the indefinite future, is questionable.

Presentation and classification. The presentation of data in the Federal Reserve *Bulletin* represents a distinct improvement over the old tables. This is true, particularly, of the functional break-down within the instalment credit sector. Other basic tables break down the amount of instalment and non-instalment credit respectively, according to holder. Here the decision to treat sales finance companies on a consolidated basis (including the operations of cash lending subsidiaries) appears unwise and should be revised. To attain the stated objective of the Board (to present the data in such detail that it can be taken apart and put together by analysts with various interests and various types of problems), much finer break-downs are necessary—and could be obtained if the Board made a serious attempt to do so.

Recent Revisions on Consumer Credit Statistics. HOMER JONES, *Federal Reserve Board.*

The Federal Reserve Board in 1953 completed a revision of the statistics regarding consumer credit outstanding from 1939 to date. The new data are the first thorough revision of the series since their inception in 1940 and were made possible by the results of the 1948 Census of Business. The revision resulted in a substantial increase of estimates of consumer instalment credit outstanding and a substantial reduction in estimates of charge accounts outstanding and a moderate increase in total consumer credit. The revised data and a description of the process of revision are presented in the Federal

Reserve *Bulletin* for April 1953. Technical aspects of the revision are discussed in a supplementary pamphlet issued by the Board. While the revision of the series was desirable and necessary, it appears that the unrevised series had never been misleading as to order of magnitude of amount outstanding or of change in outstandings. Furthermore, it is believed that the use of annual Census Bureau data from the retail trade survey will provide a much more accurate basis for estimates in the future.

In connection with the revision of these series, a thorough evaluation has been made of the concepts involved and the means of making current estimates. Denomination of the series as "short- and intermediate-term consumer credit" has emphasized the fact that it covers only part of consumer credit, since house mortgages owed by owner-occupants are excluded. The "service credit" series has been thoroughly revised, primarily by basing its major component, amounts owed to medical practitioners, upon data from the Survey of Consumer Finances. Treatment of data supplied by commercial banks has been improved upon a basis of the results of sample surveys.

The Function of the Outside Consultative Committee in the Revision of Governmental Statistics. B. D. MUDGETT.

In the revision of governmental statistics, there are five parties at interest—employers, employees, the government bureau in question, the outside public, and the whole body of statisticians interested in the particular statistical output. Employers' and employees' interests may or may not coincide with the ends sought in a revision; advantage accrues to the general public from a strong and healthy economy and statistical procedures that assist this long run objective are in the interest of the general public. The bureau personnel is in general highly competent but is beset both by routine operations that resist change and by outside pressures of groups dominated by their self interest. The whole body of scientists, as scientists, must be presumed objective and their pressures therefore lead in the direction of constant improvement of statistical processes. Herein lies the good earth that must be worked by the committee of "experts" when they are asked to sit with bureau personnel in the revision of important bodies of governmental statistics. It is this situation that gives them a function the performance of which may aid in the improvement of government statistics. They operate to perform three tasks: (1) Better theory—the clarification of concepts back of particular bodies of data or particular statistical tools through give and take discussion between consultants and bureau personnel; (2) better liaison between government bureaus and the outside public, leading to education of the public on the significance of the work of bureaus and on the competence of staff; and (3) encouragement toward working conditions that will attract competent personnel and lead to improved career service in governmental agencies

A Follow-up Study of Mortality in World War II Prisoners of War. BERNARD M. COHEN, *National Research Council*, and MAURICE Z. COOPER, *Veterans Administration*.

Using tested methods of follow-up by matching existing military, Veterans Administration, and other records, and by questionnaire survey, the Committee on Veterans Medical Problems, National Research Council, is conducting for the VA a follow-up study of mortality, morbidity, disability, and adjustment in U. S. white male Army personnel who were prisoners of war in World War II and liberated alive. Representative samples of both Pacific and European ex-prisoners, together with appropriate control groups, are being followed. This initial report is confined to mortality as observed in the first six years after liberation, and deals primarily with methods. The main methodological features are: (1) Demonstration from previous experience of the virtual completeness of VA death records of servicemen and veterans, permitting accurate measurement of mortality differentials, and providing reliable estimates of possible error where the differentials are small. (2) Selection of control samples from Army unit records to match controls with prisoners in proportion of ground and air forces, officers and enlisted men, and time distribution of and opportunity for capture, the latter two factors being better approximated in the European than in the Pacific area. (3) Use of (a) an additional control, a sample of veterans generally, to evaluate the effects on mortality of selection for military service, and (b) population life table mortalities as an intermediate device to derive stable, annual, age-specific mortality expectations.

The main findings are: (1) A marked excess of mortality in the Pacific prisoners, somewhat concentrated in the first two years after liberation, and no excess of mortality in the European prisoners. (2) In the mortality excess of Pacific prisoners, tuberculosis and automobile accidents are the most conspicuous causes of death.

Insurance Mortality Investigations of Physical Impairments. E. A. LEW, *Metropolitan Life Insurance Company*.

Life insurance mortality investigations may be regarded as a classical example of long range follow-up studies. They have usually dealt with relatively large numbers of persons who have been automatically traced for long periods of time through the circumstance of their being insured. The pro-

cedures for such studies were fully developed by actuaries more than a hundred years ago. The paper draws attention to the salient features of several types of life insurance mortality investigations and to the essential procedures used in such long range follow-up studies of physically impaired lives. It outlines the scope and the principal findings of the more important investigations of physical impairments, and discusses both the limitations and the special value of the data for public health and medicine.

The following mortality investigations of physical impairments are specifically referred to: Medico-Actuarial Mortality Investigations, 1912–1914; Medical Impairment Study, 1929 and its Supplement; Medical Impairment Study, 1936; Medical Impairment Study, 1938; Blood Pressure Study, 1939 and Its Supplement; several recent studies of the effect of build on mortality; several recent studies based on follow-ups of disabled policyholders.

Reference is also made to the Impairment Study, 1951, a comprehensive intercompany investigation of the past fifteen years' experience under some 132 groups of physical impairments, which is due to be published by the Society of Actuaries in the spring of 1954.

The findings of these investigations have made it possible to express the long range prognoses for a wide variety of impairments in numerical terms. They have been particularly valuable in shedding light on those impairments that fall in the broad range between good health and disease as recognized by clinicians, for example, overweight and moderate elevation in blood pressure. Follow-up studies of policyholders who recover after incurring a disability while insured have produced data indicative of the long range prognoses for the more serious impairments, such as advanced tuberculosis and coronary occlusion.

Factors in Interpreting Mortality After Retirement. ROBERT J. MYERS, *Social Security Administration.*

Currently there is considerable discussion as to the advantages of individuals continuing in employment beyond age 65 rather than being forced to retire compulsorily. Such advantages accrue both to the individual and to the nation. One of the subsidiary advantages frequently claimed is that an individual who is compelled to retire will lose his vitality and die much earlier than if he were allowed to continue in gainful employment. This runs contrary to the viewpoint frequently expressed several decades or more ago that workers were being kept in harness until they dropped dead from exhaustion rather than being allowed to spend their declining years in peace and leisure.

Unfortunately, specific and reliable data as to the effect of retirement on mortality are not available. The analysis is complicated by the question as to whether people retire because they are disabled and thus subject to high mortality or whether the retirement itself produces the high mortality. Clear evidence is available that retired persons do have higher mortality than active workers and the general population at the same ages, especially in the first few years after retirement. These data are presented for various Governmental retirement systems and for certain selected non-Governmental programs. Consideration is also given for various types of retirement systems as to how the resulting mortality experience may develop and what biases and limitations may be present solely because of the particular provisions of the plan rather than because of any underlying mortality effects.

Some Observations on the Inequality of Incomes. HERMAN P. MILLER

Throughout history philosophers have speculated about reasons for the inequality of incomes. Various explanations have been offered. Some have stressed ability, others chance, and still others institutional factors which give the children of the wealthy an undue advantage. This paper offers a statistical explanation for the skewness of the income curve. The thesis is that the skewed income distribution reflects the merging of several symmetrical curves which differ only with respect to the level and spread of incomes. To explain income inequality it is first necessary to understand the reasons for the differences in the component parts of the income curve.

Much of the skewness of the income curve is due to the inclusion of women in the distribution The difference between income distributions for men and women has little to do with chance, ability, or the possession of private wealth. Among men, nearly three-fourths of the highest income group (over $10,000) are independent professionals, businessmen, or managers. To the extent that there is freedom of entry into these occupations, income differences between these groups and others may merely represent the payment by society for rare skills or risk taking. The facts regarding freedom of entry are not now adequately known.

Value of Dwellings in Relation to Income. MARGARET REID, *University of Chicago.*

This analysis of value of owner-occupied dwellings in relation to income using data from the 1950 census is still in a preliminary stage. Major findings include the following: (1) Coefficients of elasticity of value of dwelling in relation to income with grouping by income of primary families and individuals occupying them had a considerable range, e.g. from .18 in Cleveland to .63 in Birmingham. For S.M.As (Standard Metropolitan Areas) in general it appears to be around .30. ($V = a + bI$, where V equals the

logarithm of value of owner-occupied dwelling and I equals the logarithm of income of primary families and individuals occupying them.) Many factors appear to contribute to this relatively low coefficient as well as to differences among S.M.A., e.g. distribution of hoseholds by age and sex of head and random variations in income. Households with head 65 years or older or with female head tend to have a high value-income ratio and tend also to concentrate at low incomes. In addition, random variations of incomes tend to concentrate at low incomes persons with a transitory decline in income and at high levels those with a transitory increase. This distribution of transitory incomes interferes seriously with differences in current income indicating what families tend to do when changes occur in income status. However the coefficient of .30 is quite similar to the corresponding coefficient of imputed rent of owner-occupied dwellings in relation to income, for places surveyed by the U. S. Department of Labor during the thirties. (2) The average value of owner-occupied dwellings for various S.M.As., standardized for income level of the family, is related to the income of the S.M.A. This is consistent with earlier observations that expenditure levels appear to be a function of the income of the community as well as the current income of families. There seems good reason to believe that if the effect of random variations in family income is eliminated the association of value of dwelling to community income as a separate factor is no longer present. (3) Coefficients of elasticity of value in relation to income from a grouping of households by value of dwelling tend to approximate 2.0. Since value of dwelling probably has less random variation than current income, it may be that this regression provides a better measure of the effect of difference in income status on the value of dwelling than does a grouping by current income. Differences in level of the value curve among S.M.As. from this classification are related to age distribution and not to income of the community. (4) Intergroup value-income relations among major portions of the New York S.M.A. and among census tracts of several places yielded coefficients of elasticity of value of dwelling to income around 2.0.

Changing Geographic Patterns of Migration in the United States. HENRY S SHRYOCK, JR., *U. S. Bureau of the Census.*

In order to study trends in net migration by States and possible changes in the overall pattern of interstate migration, figures were compiled from a variety of sources. Many of the data used were estimates especially prepared for this purpose. The following nine periods were included: 1950–1952, 1949–1950, 1945–1950, 1942–1945, 1940–1942, 1935–1940, 1930–1935, 1920–1930, and 1910–1920. Using the States as units, product-moment correlation coefficients were computed for all pairs of time periods. Of these 36 correlations, 31 were significantly different from zero at the .05 level and were all positive. The median value of all the correlation coefficients was +.58, and the range was from −.03 to +.93. There did not appear to be any systematic changes in this pattern over time; adjacent periods were not more highly correlated than nonadjacent periods. Furthermore, the pattern by States did not seem to be related to the gross volume of interstate migration of a given period. There were indications that the pattern was sensitive to economic conditions and to war conditions. The relationship was not a simple one, however.

These conditions have more obvious effect on the figures for some of the individual States. Many of these States do show fairly definite trends in net migration. The year from April 1949 to April 1950 had the most atypical pattern of interstate migration. For this year and the period 1935–1940, statistics are available on migration in both directions between pairs of States. Four States (California, Michigan, Tennessee, and Texas) were selected for intensive examination. For each of these States, the important interchanges with other States were listed, and changes between the two periods were noted. Some of the most important shifts in migratory currents were thus detected. The quality of the basic data, additional information needed, and some questions for future examination are also discussed.

The Problem of Improving Mineral Statistics. J. E. MORTON, *Cornell University.*

The problem of improving mineral statistics is only a part of a universal problem: the improvement of the systematic mass production of economic data in general. That the collection of mineral statistics should create a special problem may be attributed to the late but rapidly growing interest in the mineral sector of our economy; in part, the problem is due to the technological peculiarities of the extractive industries. Weak-spots and other "pathologies" in a given fact finding machinery are easily identified if one projects such machinery against generally accepted requirements and standards. Examples of such requirements applying to any efficient and well organized mass data production process are: (1) The proper specification of the end product; (2) The efficient production of the data while adhering to accepted statistical standards; (3) The quality control of the product to protect the data-consumer and to yield a basis for continuing improvement of the data production process. Comparing—from the above points of view—mineral statistics with other major types of economic data collection systems, with agricultural, manufacturing, population and employment statistics, one must admit that mineral statistics frequently violate the above sketched requirements. The resulting weaknesses have been reflected in such diagnostic reports as that by the Hoover Commission, the President's Material Policy Commis-

sion and the American Statistical Association's Survey of the Statistical Operations of the Bureau of Mines.

In conclusion, it appears that several basic steps will have to be undertaken before special therapeutical measures can be applied effectively. These steps are: (1) The development of a broad, yet specific, mineral statistics policy which will not only recognize the difference between the needs of the technologist on one hand and those of the business analyst and economist on the other, but also the discrepancy between the commodity and the industry-wide point of view. (2) The adherence to modern statistical standards and the development of new methods and procedures where justified by the peculiarities of the extractive industries. (3) The reconsideration of the administrative framework within which efficient production of mineral statistics is to develop, including the very important question of how to generate, attract, and best utilize the particular kind of rare combination of talents and skills needed for the successful operation of a mineral statistics program.

The Economic Outlook for 1954. GERHARD COLM, *National Planning Associates.*

To maintain "full employment" next year an increase in total production of about $10 billion would be necessary. Surveys of present spending intentions of consumers, business, and government for the next year do not indicate any sharp increases or decreases in total buying but on balance a mild further decline. The defense demand of the Federal Government may be down by $2 or $3 billion next year, but the continuing rise in state and local expenditures for roads, schools, hospitals, and other improvements will offset much of the decline in Federal demand. Business enterprises are planning only slightly less spending for plant and equipment in 1954 than in the peak year 1953. According to a recent survey of consumer attitudes, people in general are optimistic about next year's income and feel that the present is a "good time to buy." However, some decline in consumer incomes is likely. Since there is no indication that consumers intend to spend a larger proportion of their incomes, consumer spending is likely to show a moderate decline. If one simply puts the fragmentary indications of present intentions into a coherent picture and allows for some decline in inventories, one would reach the conclusion that the year 1954 would bring a reduction in total production of about $10 to $15 billion instead of the desirable full employment increase of $10 million. This would mean a level of activity of 5 or 6 per cent below the full employment level. Unemployment might rise from 1¼ million in 1953 to perhaps 3½ million in 1954.

Before accepting this as a "forecast" it must be recognized that present intentions, which are the basis for this outlook, may well be changed. In view of the possible weakness in markets for some goods, it is possible that business might engage in rather substantial inventory liquidation and might also revise downward its expansion programs. On the other hand, forward looking businessmen may think that conditions warrant a stepped up modernization program. They might also push ahead the development of new products and make the purchase of goods more attractive so that consumers are persuaded to use some of their liquid reserves for increased purchases. Finally, the Government in response to an economic downturn might adopt tax reductions beyond present plans, might adopt financial measures to stimulate the construction of residential houses, or might step up other useful programs. In our present state of knowledge it is only possible to indicate the economic trend on account of *present* intentions. This trend is mildly down. If the community maintains its confidence that the Government is ready and willing to act, it should be possible to prevent the downturn from developing into a depression. How consumers, business, and government respond to this trend, whether their responses will aggravate, mitigate, or reverse it, can be a subject of discussion but not of any forecast which pretends to be more than one man's opinion.

Economic Forecast of the Agricultural Situation, 1954. OBIS V. WELLS, *U. S. Department of Agriculture.*

No marked change in the domestic demand for food and other agricultural products appears likely in 1954 as compared with the current year. Also, foreign takings of United States farm products, while sharply reduced in the 1952–53 season from other recent years, appear to be at a level sustainable over the next year or so. Supplies of most farm products are expected to continue large in 1954. Carryover stocks may increase further by the end of the current marketing year, but a large part will be held by the Government. Acreage restrictions are likely to bring smaller wheat and cotton crops in 1954 and price support programs will continue to cushion the impact of large supplies on farm prices. With prospective conditions of demand and supply for farm products in 1954 approximately the same as in 1953, the average of prices received by farmers may hold near current levels. With cost rates to farmers stabilizing, the cost-price squeeze in agriculture is not likely to be intensified significantly in 1954.

Achieving Maximum Prediction per Unit of Testing Time. JOHN T. DAILEY, *Bureau of Naval Personnel.*

In an aptitude battery of finite length composed of pools of homogeneous items, how long and thus how reliable should each test be in order to maximize the composite validity of the battery and thus obtain maximum prediction per unit of testing time? For tests composed of homogeneous items, test validity and reliability vary concomitantly with the number of items, and varying the test length alters

both reliability and validity in a predictable manner. Formulas are derived and presented for predicting correlations with tests of altered length.

When homogeneous items are added to a given test, the amount each successive item adds to the multiple validity falls off sharply. Items added last to a long test of highly valid type items may add less than items added first to a less valid type of test. Data are presented to predict the results of a shortening of each test in the Aviation Cadet Classification Test Battery. It is demonstrated that, for a given amount of testing time, appreciably greater multiple validity may be obtained by using a large number of short tests rather than a smaller number of long tests. These predictions were verified by an empirical study.

Family Interaction and the Transmission of Achievement-related Attitudes. FRED L. STRODTBECK, *University of Chicago.*

Forty-eight recorded discussions between father, mother, and adolescent son were analyzed and significant negative correlation was found between the power of the father in the family decision-making and the score of the son on an achievement-related attitude scale. The families were second generation (the son third), of Jewish and Italian ethnicity, divided between over- and under-achieving students and stratified into three socio-economic status groups forming a $2 \times 2 \times 3$ factorial design. The frame for selecting the sample was created by administering questionnaires to all children between 14 and 17 in parochial and public schools in an Eastern City. The study is a part of the research on the early identification of talented persons sponsored by the Markle Foundation through the Social Science Research Council.

The Validation of Testing Programs for University Students. WILLIAM B. SCHRADER, *Educational Testing Service.*

The widespread use of tests for selection and guidance in American universities has been accompanied by the growth of testing programs. A testing program may include all the steps needed in collecting data on test performance by examinees and in transmitting the test results in a convenient form to appropriate test users. Validation of tests for predicting academic achievement contributes to a program by aiding in the evaluation of program components, in the improvement of test offerings, and in the effective use of test results.

A comparative study of mathematical aptitude and achievement materials for predicting engineering school grades provided useful information for an administrative decision by the College Entrance Examination Board. A simple method of assigning optimal testing times to test parts indicated that the Law School Admission Test could be appreciably shortened while maintaining the same validity. Graphs have been used to aid in combining test scores with previous academic record and expectancy tables have been developed to aid test users in interpreting predictions. The planning and interpretation of validity studies should provide for: sampling institutions according to a defined plan, using test scores jointly with other predictors, studying homogeneous student groups, avoiding capitalization on chance, taking account of restriction of range of talent, using a regression approach wherever possible, and reporting results simply. Criterion development is a promising though difficult field. Broadening of the criterion to include other major outcomes of college and professional education than those reflected in grades is needed. Also needed is long-range validation of the tests administered in the 1920's against criteria of adult success.

Statistical Principles of Testing. JOHN MANDEL

The answer to many problems in science, both fundamental and applied, is found in small differences in the numerical values of a few measurements. In determining these differences, it is necessary to guard against biases introduced by the testing procedure. One of the principal functions of statistical design is to detect and neutralise these biases. Examples are given to illustrate this approach:

(1) In an experiment made for the purpose of testing the homogeneity of 4 batches of polyisobutylene by means of flow-vicosity measurements, a 4×4 latin square was used. Considerable day-to-day variation, as well as systematic effects of chronological order within days were found. Correction of the data for these systematic effects permitted an evaluation of the degree of homogeneity of the material within prescribed limits of uncertainty, that would not have been possible without statistical design. (2) In many cases experiments run in parallel display much greater agreement than experiments run at different times or in different laboratories. This fact can be used to increase precision by the well known device of the "control sample." Statistical methodology has broadened the idea of the control sample to include the concept of the statistical "block," thereby eliminating the need for an actual control sample in many instances. The idea is illustrated by a road test of eight automobile tire brands for rate of tread wear, run in accordance with a chain-block design with two-way elimination of heterogeneity. Both run-to-run variation and the effect of wheel position are eliminated from the comparison

of the brands. (3) In interlaboratory experiments, the specimens for test are usually allocated entirely at random among the various laboratories. A judicious utilization of the idea of the experimental "block" can sometimes increase the precision several fold. Sheets of vulcanized rubber often show considerably more variability than that observed on any one sheet. In an aging study of rubber, carried out as an interlaboratory experiment, the information per measurement might well have been increased six-fold if the specimens necessary to obtain individual aging curves had been taken from the same sheet

Some Comments on the Lot Plot Plan. L. E. Moses, *Stanford University.*

The lot plot plan of sampling inspection was conceived by Shainin as a method to give a high degree of protection against acceptance of lots with fraction defective in the vicinity of 1/10 of 1%. The plan is supposed to be effective regardless of the character of the lot. It has found wide adoption in many industries in this country. Little is known of its theoretical basis. The plan calls for taking a random sample of 50 observations and plotting the histogram for the sample, as well as the mean and average range (from the 10 sets of 5 observations). From the appearance of the histogram the inspector decides whether to view the lot as "normal," "flat-topped," "long-tailed," "bimodal," "skew," or "truncated." For each such type there is a somewhat different way of deciding upon the acceptability of the lot.

This paper presents indications that the operating characteristics of the plan depend markedly on the character of the lot; that there is little hope of using a 50-observation histogram to detect "small" departures from normality which greatly distort the O C. curve. The special types of analysis prescribed for skew and bimodal samples are considered, and found to have rather unsatisfactory properties in general.

Continuous Sampling Plans. Harry Weingarten.

Continuous sampling plans (for acceptance inspection of material not assembled in "lots") available in the literature, were designed for a manufacturer interested in keeping a check on his production and also guaranteeing an AOQL. The plan proposed by H. F Dodge, "A Sampling Inspection Plan for Continuous Production" in the September 1943 issue of the *Annals of Mathematical Statistics*, is such a plan. Dodge's plan requires the alternation of 100% inspection and sampling. Two quantities i and f are specified: if after starting with 100% inspection, i consecutive items are found free of defectives, sampling at the rate f begins. If a defective is found during sampling, the inspection returns to 100% until again i consecutive items are free of defectives, etc. For a fixed value p of the incoming per cent defective, i and f determine the AOQL.

Emphasis in adopting the Dodge Plan for use by a purchaser (Bureau of Ordnance, Navy Department) was placed on reducing the amount of inspection required. If poor quality were submitted for inspection, the purchaser would, in effect, perform a screening operation for the manufacturer. The adaptation utilized the rejection number concept of lot acceptance plans; thus the Dodge Plan is operated as long as less than $a+1$ defectives are found. When inspection is interrupted, upon finding the $a+1st$ defection, it is not resumed until the manufacturer finds and removes the cause for defective product.

A sampling plan in Navord Standard 81 is defined by four quantities, N = anticipated volume of production, i = length of 100% inspection requirement, f = sampling rate, and a = maximum number of defectives allowed before interruption. This generalization of the Dodge Plan (Dodge Plan: $N=\infty$, $a=\infty$) produces mathematical problems of great complexity. If the effect of a plan defined by N, i, f, and a is described by

$$L_p = Pr \text{ (not interrupting inspection)},$$

it was possible to find L_p exactly, only for $a=0$, 1, 2. For larger values of a an approximation was used.

Problems of Coordinating the United States Statistical System. Stuart A. Rice.

The statistical system of the United States embraces many official, semi-official, and unofficial agencies and instruments, among which some degree of coherence is maintained by item to item adjustments among related tasks and processes. The relationships may be those of conceptual congruity or those of consistency in operational patterns and sequences. The process of maintaining coherence is called coordination. In the decentralized statistical system of the United States a particular problem which falls to the Office of Statistical Standards as the Government's central agency of statistical coordination is that of representing the general interests. These often override or lie between the interests and responsibilities of particular statistical agencies. Although a cross-cutting "statistical budget" is developed by OSS each year, its components must be fitted into the over-all budgets of departments and agencies, thus raising special problems. Other important problems of coordination are the establishment of balance among separate agency programs, of establishing lines of demarcation between governmental and non-governmental responsibilities and of finding the appropriate boundaries to the conceptions of "confidentiality" and national security interest in data. Finally, the problem of liaison between Federal

statistical agencies and the statistical profession is being solved in piecemeal and modest but satisfactory ways through advisory mechanisms established by the American Statistical Association.

Problems of Co-ordination in the Canadian Statistical System. HERBERT MARSHALL.

Canada's statistical system is highly centralized. The Dominion Bureau of Statistics was set up under a Statistics Act in 1918 and instructed to "organize a general scheme of co-ordinated social and economic statistics pertaining to the whole of Canada and to each of the provinces thereof." To reach this objective several types of co-ordination were necessary. Since Canada is a federal state the raw materials for some kinds of statistics are derived from the administrative records of both federal and provincial departments. These administrative records have to be fitted into the general overall scheme. Co-ordination with provincial departments is achieved mainly by annual or less frequent Dominion-Provincial Conferences in various statistical fields. The bulk of the Bureau's output is based on information collected directly rather than from administrative records and includes censuses of population, agriculture, industry, distribution, labour and prices, and numerous other fields. Co-ordination with those who fill in the questionnaires and those who use the data involves close liaison with numerous business and other organizations and a wide variety of users of the data. Co-ordination of the work done within the Bureau of the fourteen Divisions and numerous sections is of great importance. Uniformity in concepts, definitions, classifications, avoidance of duplication, and other aspects of co-ordination require constant vigilance on the part of interdivisional committees, and other supervisory efforts.

Methodological Problems and Findings of Study of Recipients of Old-Age Assistance. THOMAS G HUTTON, *Department of Health, Education, and Welfare.*

The major methodological problem faced by the Bureau of Public Assistance in making the study of old-age assistance recipients, as in all its studies, is how to collect data that are comparable from State to State when the Bureau must conduct its studies with personnel of 53 State public assistance agencies each operating under its own, and frequently diverse, policies, and with case workers in over 1,000 governmental units scattered throughout the length and breadth of the land. These workers have diverse backgrounds and varying degrees of technical skill. Contacts with State personnel and local case workers are almost entirely through the written word, and the Bureau is dependent on obtaining the cooperation of these workers in order to obtain data that are reasonably accurate. The solution to this problem while not a statistical one, but one in the field of human relations, is an integral part of making such a study. The paper discusses how this problem was solved.

The study of old-age assistance recipients was designed to obtain answers to two questions: (1) Who are these old people that are receiving public assistance, and (2) How do they live. Answers to the first question were obtained in terms of certain social characteristics; age, sex, race, marital status, and physical and mental condition. Answers to the second question were obtained through information on living arrangements, extent of home ownership, certain housing characteristics relating to extent of overcrowding, sanitary facilities in the home, and the use of modern conveniences such as electricity, the telephone and refrigeration. The degree to which children contributed to the support of aged parents, the amount and sources of income other than assistance, and the total amounts of income on which these people lived also contributed to answering the second question. The findings of the study on each of these points are presented in brief compass in the paper.

Methodological Problems and Findings of Survey of Aged Beneficiaries of Old-Age and Survivors Insurance. EDNA C. WENTWORTH, *Social Security Administration*

In the fall of 1951 the Bureau of Old-Age and Survivors Insurance conducted a Nation-wide survey of the economic resources of retired worker and aged-widow beneficiaries. The paper discusses the methods used in evaluating the economic situation of the beneficiaries and some of the findings for those who received benefits throughout the survey year. Slightly over three fifths of the beneficiaries had income in addition to their own independent retirement income, or they used assets. The chief sources of the additional income were public assistance, contributions from relatives outside the household, and earnings, usually from short-time employment. Two thirds of the beneficiaries either received public assistance or had less money income from all sources than public assistance would allow its recipients who lived alone in rented quarters; a sixth were assistance recipients. Half of those with the low incomes and no public assistance got along because they lived with relatives and were partially supported by them. Some of the others had noncash income of various kinds. Altogether, half the beneficiaries were partially dependent during the survey year, more of them on relatives than on public assistance.

If they used one tenth of their liquid assets each year, most (70 per cent) of the retired workers with the highest benefits—$60 to $68.50—would have independent retirement funds for the next 10 years of at least $900 a year if single and $1,500 if married. Only 36 per cent with benefits of $50 to $59 would

have such independent retirement funds. Those with smaller benefits were worse off. Homes were owned by 45 per cent of the beneficiaries but the homes are not taken into consideration in this appraisal of beneficiary resources.

Intervals Between Onsets of Multiple Cases of Poliomyelitis in Families. ARTHUR S. LITTELL and GEORGE V. SMITH, *Western Reserve University.*

Recently P. E. Sartwell has shown that observed incubation periods of poliomyelitis are described by a log-normal frequency distribution. Using this distribution of incubation periods an expected distribution of intervals between onsets of pairs of cases exposed simultaneously was computed. Six observed series of intervals between onsets of initial cases and onsets of subsequent cases within families, 1310 intervals in all, are compared with the distribution expected if all cases in each family resulted from common exposure. The agreement between observed and expected distributions is very good in some of the series. In the others, the disagreements are opposite to what one would expect if there were an appreciable proportion of secondary infections within families. Instead of an excess of intervals longer than 7 days, which would represent secondary cases, there is a deficiency.

It is shown that, even if non-susceptible carriers are partially responsible for the spread of poliomyelitis, an observed distribution of intervals should show an excess due to secondary cases of intervals longer than 7 days over the number expected on the basis of common exposure. This argument points to one of two conclusions: (a) when multiple cases of poliomyelitis appear in a family, the source of infection is common, and spread of the disease from case to susceptible within the family is rare; or, (b) a ase of poliomyelitis can be infectious very early in its incubation period.

Unsolved Problems of Experimental Statistics. JOHN W. TUKEY, *Princeton University.*

It would not be misleading to say that there is only one unsolved problem of experimental statistics —"How can we identify the problems of experimental statistics?" (We can identify a good many unsolved problems by accident, but we probably miss many important ones for far too many years.) Experience to date indicates that difficulties in identifying problems have delayed statistics far more than difficulties in solving problems. This seems likely to be the case in the future, too.

Thus it is appropriate to be as systematic as we can about unsolved problems. Any system may be a start toward a partial solution of this one central unsolved problem. We shall try to do this by stating first some hypergeneral principles and then some general consequences. We shall strive to phrase these as generally as possible, in the hope of prolonging their useful life. The discussion of examples of these 18 general principles will set forth a number of unsolved problems, while a list of 37 provocative questions poses many more. The account closes with a discussion of the possibility of orienting experimental statistics toward problems rather than techniques.

Some general principles. If we feel that the detailed problems of experimental statistics arise from the interaction of certain general principles among themselves and with classes of experiments, it is reasonable to try to state and illustrate some of these principles. Most of these hang on four hypergeneral principles, which may seem harmless until we come to their consequences, namely: (A) Different ends require different means and different logical structures. (B) In each area, statistical method must and does evolve, mainly by adding *both* immediate ends *and* considerations. (C) While techniques are important in experimental statistics; when to use them and why to use them are more important. (D) In the long run, it does not pay a statistician to fool either himself or his clients.

Characterization of Distribution-free Statistics. Z. W. BIRNBAUM.

1. *Definitions.* Let Ω and Ω' be families of cumulative probability functions. A real quantity $W = S(X_1, X_2, \cdots, X_n, G)$ is a *statistic in* Ω *with regard to* Ω' if, for any $G \epsilon \Omega$ and $X_1, X_2, \cdots, X_n$ in the n-dimensional sample space E_n of a random variable X with the c.p.f. $F \epsilon \Omega'$, this quantity 1^0 is defined almost everywhere in E_n and 2^0 has a probability distribution; this probability distribution is then denoted by $\mathcal{P}[S(X_1, \cdots, X_n, G); F]$.

If $\Omega = \Omega'$ and the statistic $S(X_1, X_2, \cdots, X_n, G)$ has the property that $\mathcal{P}[S(X_1, \cdots, X_n, G); G]$ is independent of G for $G \epsilon \Omega$, then $S(X_1, \cdots, X_n, G)$ is a *distribution-free statistic in* Ω.

Let Ω be a family of c.p.f.'s such that the inverse function $G^{(-1)}$ can be defined for each $G \epsilon \Omega$. A statistic $S(X_1, \cdots, X_n, G)$ in Ω with regard to a family Ω' is called *strongly distribution-free* if $\mathcal{P}[S, X_1, \cdots, X_n, G); F]$ depends only on $\tau = FG^{(-1)}$ for all $G \epsilon \Omega$, $F \epsilon \Omega'$.

If there exists a function Φ defined on the n-dimensional unit cube and symmetric in its arguments such that for any $G \epsilon \Omega$, $F \epsilon \Omega'$ we have $S(X_1, \cdots, X_n, G) = \Phi[G(X_1), \cdots, G(X_n)]$ almost everywhere in E_n for the random variable X which has c.p.f. F, then $S(X_1, \cdots, X_n, G)$ is a *statistic of structure* (d) in Ω with regard to Ω'.

2. *Problems.* Some types of problems which may be formulated in terms of the concepts defined above are: (a) given Ω, characterize the class of statistics which are distribution-free in Ω, or strongly dis-

tribution-free in Ω with regard to a given Ω'; (b) given Ω and a loss-function L, determine the distribution-free statistic which is optimum for a specified decision-theoretical problem.

The following are some results obtained by the speaker and H. Rubin for problems of type (a): If a statistic in the class Ω_1 of all continuous c.p.f.'s with regard to Ω_2 has structure (d) then it is distribution-free in Ω_2; but not every statistic of structure (d) in Ω_2 with regard to Ω_2 is distribution-free in Ω_2. A statistic in the class Ω^* of strictly increasing continuous c.p.f.'s with regard to Ω^* is strongly distribution-free if and only if it is of structure (d). For problems of type (b) O. P. Aggarwall has recently shown that for estimating a cumulative probability function the minimax procedure invariant under all permutations of the sample is, for certain plausible loss functions, either that based on Kolmogorov's statistic or on some modifications of that statistic.

Inadequacies of the Construction Estimates as General Economic Measures. WALTER E. HOADLEY, JR., *Armstrong Cork Company.*

Construction merits general recognition as a major American industry, but few leaders within this diversified industry show much interest in aggregate activity statistics, and most "outsiders" have so little first-hand knowledge of construction that they tend to accept almost without question any available measures of construction, especially from "official" sources.

At least five general inadequacies stand out in the construction statistics field at present: (1) a widespread lack of understanding and appreciation of the importance of adequate construction statistics to public and private policies, (2) weaknesses in the basic construction series currently available, (3) major gaps in construction statistics, (4) the absence of a comprehensive plan to strengthen government statistics in the construction field, and (5) insufficient funds to insure suggested improvements can be carried out successfully. As construction activity has grown in size and importance to the whole economy, our knowledge about it has lessened. This unfortunate condition is not due to lack of competence in the governmental statistical agencies, which are doing just about the best job that can be done with the resources available. It is due primarily to the failure of Congressmen, businessmen, and others to recognize the importance of sound data as a basis for market planning or to recognize just how inadequate the present data are for this purpose.

Since the Korean War, for the first time a construction estimate series—the nonfarm housing start series of the U. S. Bureau of Labor Statistics—was written into law as a "trigger" statistic to help time changes in a government control (i.e., residential credit). This last move, perhaps more than any other single development, stimulated new interest in government statistics across the home building industry. Government construction specialists also quickly recognized the need to reexamine the adequacy of their statistics in light of their increased role in policy determination. Nevertheless, few major policy decisions affecting the construction industry can now be properly made on the basis of the statistics available.

The most encouraging development this year has been the extent to which policy level people within and outside government have become aware of many deficiencies in construction statistics and have started to promote plans designed to improve several key series.

The three principal statistical gaps are: (1) Current indications of changes in the national housing inventory, reflecting the continuing impact of new buildings, conversions, and demolitions. (2) Estimated vacancies in old and new dwellings. (3) Estimated expenditures for "fix-up" purposes (i.e., repairs, additions, and alterations to existing construction). These data are absolutely essential for the housing and construction segment of the economy to prosper and avoid major swings.

The pressing job to be done in government—with the aid of all persons interested in construction—must be to obtain some general agreement as to the basic objectives of government in the construction field. Government policy officials must decide what *they* feel *they* need to make sound decisions in the interest of growth and stability in the economy. Only when the major deficiencies and gaps in "public policy" statistics on construction have been substantially eliminated should attention be directed to meeting the numerous and ever-present requests for more detailed market type information—except of course as such becomes available as a by-product of public policy statistics. In my view, the construction industry and the economy as a whole have much more to lose from inappropriate public policies than from the absence of fairly detailed information about particular segments of construction markets.

There is reason to believe that construction statistics programs will be strengthened in the months ahead—not simply because of our wishes, but rather because "public policy" construction statistics—on their own merits—promise to be recognised more widely as far too important to public policy determination to be allowed to remain inadequate or to deteriorate further.

Analysis of Variance Models in Sedimentary Petrology. J. C GRIFFITHS.

Investigation of the petrography of sedimentary rocks by analysis of variance already embraces completely randomised, randomised blocks, unbalanced lattice and factorial designs. The hierarchy of petrographic units represented by rock-types, formations within rock-type, outcrops within formations,

specimens within outcrops and aliquots within specimens leads to multistage sampling patterns and completely randomized designs.

Interacting factors and randomized blocks design may be introduced by using two or more spatial dimensions analogous to "rows and columns" of agronomic experiments. In contrast, however, the aim of such arrangements in the analysis of sediments is the evaluation, and not solely the segregation, of this "fertility gradient". As another interacting factor order may be introduced either as a stratigraphic variable or as a sequential arrangement of a series of experiments. Equivalent to the "variety" effect in agronomic investigations, observers performing the same experiments, leads to evaluation of operator variation. In this connection operator inconsistency or "discrepance" has proved of vital interest in setting the sensitivity level of many petrographic investigations.

Heterogeneous variance has complicated many of the comparisons but by estimating components of variation it has been shown that mean differences are, perhaps, less important than differences in variability in characterizing sedimentary rocks and the processes by which they are formed. Equal sampling based on the assumption of equivalence of classes, as for example in using, arkose, graywacke and quartzite, has proved fallacious and introduces heterogeneous variance. Sampling in proportion to variability appears to be the obvious solution as emphasized by some recent analyses of variance of sphericity and roundness of quarts grains in sediments.

On Testing the Association of Mineral Occurrence with a Set of Observable Characteristics. RAY MICKEY.

Let a topographic map M be given for an area to be explored for the presence of a set of minerals. Let V be the vector function defined over M which associates with each point of M a set of values of geological characteristics, and let Z be the function over M which assumes the value one for those points whose mineral density is sufficiently large and which assumes the value zero otherwise. Let x be a random point in M whose distribution has constant density over M. We will say that V and Z are associated if $V(x)$ and $Z(x)$ are not independent chance variables.

We consider the possibility of testing H_0, the hypothesis of no association, by the use of conditional tests of the following class. Let $t = t(V_1, \cdots, V_n, Z_1, \cdots, Z_n)$ be a test statistic with observed value t_0. Conduct a series of sampling experiments in which $V_1, \cdots, V_n$ and $\sum Z_i$ are held fixed and H_0 is assumed. The test is conducted by the use of a binomial sampling plan. Perhaps the simplest of these of level $1/k$ are: reject H_0 if

$$\sum_{i=1}^{mk-1} u_i \leq m - 1,$$

where $u_i = 0$ if the value of t obtained from the ith sampling experiment is less than t_0 and $u_i = 1$ otherwise

An example is presented in which use is made of the statistic

$$t = \max_{i \in S_\alpha} \min_{j \neq i} \rho(V_i, V_j)$$

where S_α is the set of indices for which $Z_i = \alpha$, and ρ is a suitable metric.

Problems in Sampling the Phosphoria Formation. R. A. GULBRANDSEN and V. E. MCKELVEY, *U. S. Geological Survey*

The Phosphoria formation of Permian age contains vast resources of phosphate and other materials distributed over an area of 135,000 square miles in eastern Idaho and adjacent states. In an effort to appraise the total mineral resources in the formation and to determine its origin, the U. S. Geological Survey has smpled the full thickness of the phosphatic members of the formation at about 200 scattered localities. The sampling plan was based mainly on geologic intuition, or geologically educated judgment; but it took account of previous information available on the variation in composition and thickness of the beds, the distribution of the rocks and their outcrops, and the purposes the samples were expected to fulfill.

All samples collected are being analyzed for phosphate and other constituents of prime economic importance. It is expected that the average of five to fifteen widely spaced samples collected from an area of a few tens of square miles will represent the grade and thickness of deposits in that area and that the range in grade and thickness in any one area will indicate the magnitude of variation to be expected there. The chief uncertainty in estimates of reserves will result from lack of information on the presence or absence of beds at depth. Some of the samples will be analyzed for constituents or properties of scientific interest only. These samples will be selected randomly from samples grouped according to rock-type, stratigraphic position, or area.

Longitudinal Studies of HIP Experience, Background and Some Findings of Pilot Study. NEVA R. DEARDORFF. *Health Insurance Plan of Greater New York.*

The longitudinal studies now in progress at the Health Insurance Plan of Greater New York are essentially examinations of the experience of people covered for comprehensive prepaid medical care for periods of time up to a maximum of four years. Three phases of the experience of these people are being examined: their enrollment experience, their utilization, and the diagnoses reported for them. The happenings to a constant group of people can be reported from year to year and the changes noted in the *group's* pattern of experience, irrespective of the sequence of changes experienced by any one person or family. Another focus of interest is the *movement* of persons and families from one category to another over the years.

Longitudinal studies point toward the treatment of each person's *history* rather than his condition at a given time, as the unit of statistical manipulation. In a sense, it is the statistical approach to a group of persons or families, each with a continuing life history to be considered, rather than the construction of cross-section views of these people's characteristics made from time to time. This Project was planned and is being carried out under the direction of a committee composed of persons who have no affiliations with HIP, implemented by a specially employed staff. The work of the Project is financed by grants from the Commonwealth Fund and the Rockefeller Foundation.

Among the requirements for effective longitudinal studies is the need for a comparatively large number of subjects since almost all groups are subject to attrition from death and other causes of loss of subjects and since patterns of change tend to proliferate with each added year of observation. (A diagram illustrating these points was presented.) The HIP studies are based on random samples comprising 8,025 persons who were under observation from 1948 to the end of the study period (December 31, 1951), 9,325 since 1949, and 3,500 since 1950. Besides these, there are 8,000 persons whose coverage was terminated or had been interrupted or who were spouses and children acquired in the last year (1951) of the study period. These 8,000 cases have been included not only because of their place in annual tables, but also for purposes of some special analyses on enrollment, utilization, and diagnostic conditions. Altogether a total of 28,850 persons are included in these studies. Three tables from a pilot study were distributed to illustrate a few types of analysis which are being employed in the presentation of data.

Longitudinal Study of Health Insurance Plan of Greater New York. NATHAN GOLDFARB, *Health Insurance Plan of Greater New York.*

Although there is an increasing recognition of the particular advantages of longitudinal data as contrasted to cross-sectional data, there has been only limited experience thus far in the organization and processing of large masses of data for longitudinal analysis. It would, therefore, be extremely useful for those who are interested in longitudinal analysis to discuss the problems and methods of organizing and analyzing longitudinal data.

There are two basic sources of data for this Special Research Project, namely, (1) the documents completed by the physician and submitted to HIP at regular intervals, which identify the individuals, and give the tentative diagnosis and other related information, and (2) the registrar records which provide the enrollment data and the personal identifying characteristics for each individual. In order to determine whether the physicians' reports were being prepared with sufficient care for the needs of the Research Project, the data from these reports were checked for correspondence with the records in the physicians' files and found satisfactory. The 10 per cent sample of families and individuals covered by the HIP for which all the Project data are being obtained, was also tested and found representative for a number of characteristics.

The objectives of the Project were framed in a series of questions. It was evident that a history and summary of the medical experience for each individual for the entire period of the study was required. The Project then set about to summarize data covering approximately 30,000 people, with about 315,000 medical services, for a period ranging up to four years. Coding and tabulating methods were evolved which permitted an examination of the medical experience and care for each person in the sample. Codes were developed to give the detailed diagnoses within each study year, cumulative diagnostic codes to summarize the total diagnostic experience for an entire year, and still another set of cumulative codes to summarize the diagnostic experience for combinations of successive years of coverage—2 years, 3 years, and 4 years. Chronic diseases were classified separately from the non-chronic diseases. Every medical condition was coded, but there was no segregation of "primary" and "secondary" diagnoses, nor was there any count of episodes. These decisions were in part dictated by the nature of the reporting. Where a diagnostic description was evolved over a period of time, the "best" diagnosis was chosen on the basis of a set of rules which, for example, gave priority to the specialist's report over that of the general physician, and to the later diagnosis over an earlier one in a sequence of visits to the doctor.

The tabulations on utilization will show the trend for various cohorts of individuals indicating how, with the passage of time, the factors of age, sex or a diagnostic type affect the quantity of medical care.

"Spot Checks" in Lieu of Complete Censuses. HOWARD C. GRIEVES, *Bureau of Census.*

In the summer of 1953 the Congress indefinitely postponed the Quinquennial Censuses of Manufactures, Mineral Industries, Business and Transportation. It substituted the sum of $1.5 million for "spot checks" in the fields of manufactures, business and agriculture. The appropriation is being used for three broad purposes: (1) to fill some of the gaps in national aggregates created by the deferral of the complete censuses; (2) to increase the timeliness of reports; and (3) to develop some new statistical indicators.

Specifically, sample surveys will provide national totals of wholesale, retail, and service trades for the year 1953. The Annual Survey of Manufactures for 1953 will also be carried out. Monthly estimates of retail sales and the labor force are being improved by increasing the number of primary sampling areas. The monthly wholesale series will henceforth utilize a probability sample developed for the 1953 Annual Survey. The publication *County Business Patterns* is being compiled on a greatly accelerated time schedule covering 1953. Advance estimates of retail sales by major kinds of business employing a probability sample are being released on the 10th day following the month covered. A complete census of retail trade, using an abbreviated report form, is being taken in Dallas on a trial basis with financial assistance from local business. Finally, an effort is being made to improve the monthly measures of retail inventories. A number of other smaller projects are also in progress.

While much useful data will be forthcoming it is important to note some of the purposes of complete censuses which cannot be achieved with the resources available. The most important omission will be statistics for small areas, cities, counties, etc. Similarly, detailed characteristics, fine industry breakdowns, commodity statistics, etc., cannot be provided. The planned improvement of the samples in current use based on enumerations of the entire universe correctly classified must be deferred. Similarly, the planned coordination of Federal statistical series to achieve uniform classification is postponed until the complete Censuses are taken.

Developments in Collection and Processing of Mass Statistical Data. MORRIS H. HANSEN and WILLIAM N. HURWITZ.

Large-scale census and survey operations offer numerous opportunities for the application of statistical and other scientific methods. Such applications are now frequently referred to as operations research. The general aim in developing and improving methods is to produce the needed results at minimum cost, and within specified time schedules and other restrictions. It involves making the fullest possible use of available resources.

Modern sampling methods make it possible to provide timely results of known precision at low cost through monthly, quarterly, annual, or special sample surveys of the population, retail trade, manufacturing, and other subjects. Also, sampling methods are applied in the major censuses to collect and tabulate much of the basic information at reduced cost, to produce more timely results, to check on the quality of the census results, and in other ways.

Important uses of statistical methods in connection with both censuses and current work is in the control of quality. Quality control usually consists of a system of sample inspection and process control introduced to insure at low cost that the final product will meet certain quality standards. Quality check operations are to be distinguished from quality control. Quality checks have been introduced in situations where it has not been feasible to control quality. They involve the use of intensive measurement methods on small samples, and have been used especially in evaluating the quality of the field work in major census operations. Not only do they guide in the proper use of the results, but they provide the basis for improving the design of future censuses and surveys.

Other areas for the application of statistical methods include evaluation of alternative procedures through properly designed experiments, such as alternative types of questionnaires, collection processes, or training methods. Also, the effects of editing and other operations have been investigated and modified or mechanical techniques substituted. Our work has indicated that much editing formerly regarded as necessary, can be eliminated or reduced with only a trivial effect on the accuracy of results.

Another area for application of statistical methods is in setting standards or incentives for personnel performance. Very substantial production gains have been achieved simply by administrative incentives through performance standards set on the basis of appropriate samples of work and sample observations of workers.

Large scale electronic computing equipment is being applied and will have a rapidly increasing role in the future, as will, also mark sensing or reading methods. The introduction of such equipment, together with sampling, quality control, techniques for measuring and controlling response errors, and other methods mentioned should result in substantially increasing the timeliness of major census results

as well as reduce their costs. The application of statistical methods in the census has been the principal work of a small group of versatile scientists interested in applications. They are trained in mathematical statistics, survey methods, psychology, and other techniques, and are given complete authority to ask questions and investigate and recommend, but are kept free of operating and responsibilities.

The Current Statistics Program of the Census Bureau. A. Ross Eckler and Conrad Taeuber, *Bureau of the Census.*

Much of the current statistics program of the Bureau of the Census exists to serve needs not met by the censuses, nor by reports which are a byproduct of government administration nor by reports based on special-purpose collections by government or private agencies. Some surveys and sample censuses develop national or regional estimates for items covered in the censuses; some surveys provide essentially complete coverage for a segment of the universe covered by a census; and some current reports, like those on foreign trade, are based upon documents connected with government administrative programs. Sample surveys are of major importance in meeting the needs for current data, and in relieving the pressure for inclusion of items in major censuses.

With allowance for increases in basic costs, the resources for current statistics programs in the Census have been reduced by more than 50 percent since 1947. Although program losses have been numerous, efforts have been made to retain key items and to apply improved methods, such as the introduction of a sample of 230 primary sampling units for the current field surveys in place of the 68 area sample used previously. The Bureau's resources for collecting and tabulating statistical data are used not only for carrying out the Bureau's own program, but also in performing similar tasks for other governmental units and for private agencies, when proper safeguards concerning the public use of the results can be assured. Through the increasing use of data available from administrative records and through extension of services, where appropriate, to other governmental, as well as to non-governmental, agencies the Bureau of the Census may be successful in offsetting some of the losses in statistical information resulting from program reductions in recent years.

Preliminary Tests and Pool Rules. T. A. Bancroft, *Iowa State College.*

Certain applications in experimental design and regression analyses involving preliminary tests of significance can be arranged as special cases of a test of a general linear hypothesis in canonical form incorporating a preliminary test of significance. In order that one may have confidence in the use of such analyses it is necessary to investigate possible biases in estimates and the power and size of such complex tests. Bechhofer, using exact methods, was able to evaluate the integrals involved in the expression for the power and size of such tests in closed form for only a few special cases of the parameters involved. In order that information on power and size of such tests be explicitly available, certain approximate formulas have been developed by the author by making use of a central chi-square approximation to the non-central chi-square distribution. Empirical studies indicate that the approximate formulas give usable results, in particular for important values of parameters which have been so far unattainable by exact methods. Investigation of exact error bounds are in progress.

An Extension of Preliminary Tests for Pooling Data. D. V. Huntsberger, *Iowa State College.*

Mosteller (*Jour. American Stat. Assn.*, Vol. 43, 1948) has shown that, if $\bar{X}_1$ and $\bar{X}_2$ are independent estimates for the means μ_1 and μ_2 of two normal populations with the same known variance, the "sometimes pool" estimator for μ_1 based on a preliminary test of the hypothesis $\mu_1 = \mu_2$ is subject to large losses of efficiency, relative to $\bar{X}_1$, for some range of $\gamma = \mu_1 - \mu_2$. In this paper it is shown that, if $T = (\bar{X}_1 - \bar{X}_2)\sqrt{n}/\sqrt{2}\sigma$ is used for the preliminary test, the estimator $W(T) = \phi(T)X_1 + (1-\phi)(\bar{X}_1 + \bar{X}_2)/2$, where $\phi(T)$ is a continuous function of T, provides a marked reduction in the maximum possible loss of efficiency and extends the range $|\gamma| \leq C$ for which a gain may be realized. If $\phi(T) = 1$ for $|T| \geq T_\alpha$ and $\phi(T) = 0$ for $|T| < T_\alpha$, where Prob. $(|T| \geq T_\alpha) = \alpha$, $W(T)$ reduces to the estimator based on a preliminary test of size α. The efficiencies of $W(T)$ and the "sometimes pool" procedure are compared for a particular function $\phi(T)$. The results obtained for this example suggest that whenever a preliminary test for a possible pooling is indicated an estimator analogous to $W(T)$ will provide a substantial reduction in the maximum possible loss of power or efficiency of subsequent inferences.

Simultaneous Test of Linear Hypotheses by Analysis of Variance Methods. M. N. Ghosh, *University of North Carolina.*

In the analysis of variance for survey type of data, the tests for different hypotheses corresponding to sets of parameters are not usually independent. A notion of quasi-independence has been developed in this paper which depends upon the irrelevance of extraneous parameters in the first and second kinds of error for the test of each of the hypotheses. In the simultaneous test of the hypotheses, $H_1, \cdots H_s$, decisions of the type accept H_1 or reject H_1, accept H_2 or reject H_2, etc., are considered and

the significance level of such a test is defined as the probability of rejecting at least one of the hypotheses $H_1, \cdots H_s$, when all of them are, as a matter of fact, true. A method has been given for simultaneous test at a significance level α, when the least squares estimates of the sets of parameters for different hypotheses are non-orthogonal. These methods are applied to analyze growth data for children especially in relation to stage of sexual maturity and certain blood chemicals.

Multivariate Analysis of Covariance for a Latin Square. WILLIAM J. MOONAN, *University of Minnesota.*

It is known that, for some problems which are appropriately evaluated by analysis of variance and covariance methods, the procedures of univariate analysis must be generalized to the multivariate case. The reasons why such generalizations are necessary have been made quite explicit by H. Hotelling (1947) and others.

The multivariate analysis of variance and covariance of a 4×4 Latin Square is derived in this paper. The method of derivation follows the usual least-squares approach. However, the technique is followed for a matrix variate instead of a single variate. In particular, the multivariate normal model is assumed, the estimation space is specified, the normal equations defined, and the estimates of the parameters are made. Hypotheses on the multivariate parameters are given and the least-squares process is repeated to determine the sums of squares and cross-products which are appropriate for testing the hypotheses. A numerical example is carried through to illustrate all of the calculations necessary to evaluate the parameters and make tests of significance. The method of analysis is quite similar to the usual univariate calculations.

Weighting Coefficients for Age-Adjusted Death Rates. HARRY SMITH, JR. *University of North Carolina.*

The problem of finding a descriptive mortality index for a community is reviewed. The methods for adjusting death rates for age are discussed in terms of the criteria for the use of each. The advantages and disadvantages of each are also listed. A new method of adjusting death rates for age is proposed when the comparison of two communities is to be made. The assumption upon which this new method is based is that the mortality functions of the two communities are of the same shape and separated by a constant differential. Any overlapping that occurs is assumed due to sampling fluctuations. The solution is based upon the method of maximum likelihood. A comparison of the results using this new method with results of other methods is shown. The use of a discriminant function in determining those age groups most influential in discriminating between states in the past ten years is also studied as another possible means of calculating an over-all mortality rate.

Some Statistical Evidence on Economies of Scale. FREDERICK T. MOORE, *Bureau of Mines.*

Statistical evidence on economies of scale in manufacturing is scarce because of the necessity of compiling detailed cost studies on plants, or, on an engineering basis, computing production functions. In either case the procedure is tedious. Engineers have derived certain rules of thumb for evaluating the relationship between capital costs and capacity. One such rule is the ".6 factor" rule which states that the increase in capital cost is given by the increase in capacity raised to the .6 power. For example, the surface area of a spherical tank (which essentially measures the cost) increases as the volume of the tank (capacity) to the two-thirds power. The ".6 factor" has been applied both to process equipment and to complete plants. It is apt to fit best industries which are (a) capital intensive; (b) continuous flow operation; and (c) with a homogeneous standardized product. The chemical and mineral processing industries in general meet these criteria.

Tests were made on samples of plants from several industries, using the formula: $\log E = \log a + b \log C$ where E = capital cost; C = annual capacity and a and b are constants. Values of $b < 1$ indicate economies in capital cost and since operating costs tend to behave in the same fashion, this also indicates economies in scale. In the tests the following values for b were obtained: alumina .95; aluminum reduction .93; aluminum rolling .88; aluminum extrusions 1.0; cement .77; tonnage oxygen .63. Tests were also made for individual processes in these industries. A separate study of the production function in petroleum pipelines indicated economies of scale up to approximately 200,000 barrels of throughput per day and constant returns for larger throughputs.

It was suggested that a simplified method for evaluating economies of scale can be visualized in three steps: (1) extensive engineering information is available showing the cost of process equipment relative to some engineering magnitude (e.g., square feet of heating surface); (2) the engineering magnitude can be related to capacity by an appropriate formula (e.g., the capacity of a tank can be related to its diameter); therefore (3) by combining steps (1) and (2) above, cost can be related to capacity. This method can be applied to process equipment and the processes combined to show a complete plant. Since the method may be used for a sample as small as two, it appears to offer a simplified approach to evaluating economies of scale

Stability of Technical Coefficients: Evidence from Inter-Plant Differences in Labor and Materials Productivity. A. PHILLIPS.

Technical coefficients for Leontief matrices appear to be quite stable through time. This may arise from two distinct conditions. First, there may be a relatively stable industry technology which causes each plant to have approximately the same coefficient. Random influences affect the plant coefficients but tend to offset each other, resulting in stable aggregate industry coefficients through time. On the other hand, industry coefficients may be stable even with very different plant coefficients so long as the plant coefficients are themselves quite stable and the plants produce consistent shares of the total industry output.

Evidence from one industry shows the latter to be the case. Little central tendency of plant coefficients is found. But each plant tends to have the same inputs of labor and materials per unit output each year and produce the same relative share of industry output. Differences among plant coefficients appear related to other factors such as principal product type, integration and kind of material consumed, though most differences cannot be explained with available information. Since each plant may have its own technology which is properly part of the industry, technology is not ruled out as the cause of stability, though one fairly unique and stable industry technology does not appear to be the cause.

Adequacy of International Trade Statistics for Economic and Business Analysis. J. EDWARD ELY.

There are important conflicts between the various users of international trade statistics as to what the statistics should and can do. These conflicts have been resolved in different ways in different countries with the result that countries follow widely varying definitions and practices in compiling their foreign trade statistics. It follows from this that what may appear from the comparison of the statistics of trading partners to be errors in the statistics, are frequently differences caused by differences in definitions.

International trade statistics face a unique problem in the fact that, in contrast to all other types of statistics, more than one sovereign country compiles detailed information on the same phenomenon, in this case the movement of goods from one country to another. The detailed figures presented by a country in regard to its foreign trade are, therefore, subject to searching comparison with similar detailed figures released by another country on the same transactions. Adequate trade statistics for such multiple-country use will face major conflicts of interest with domestic uses. Since many of these and other conflicts may turn out to be irreconcilable it may be that there will have to be a substantial increase in the use of dual compilation procedures designed to fulfill conflicting needs for information.

Accuracy of Foreign Trade Statistics. OSCAR MORGENSTERN.

Accuracy of Foreign Trade Statistics can be estimated by comparing records for the same transaction by exporting and importing countries. This reveals for all periods very large differences in pairwise trade both for values and quantities. Difficulties in classification of commodities, transportation costs, tariffs can explain only part of these differences. Even if successfully enumerated, the various factors work jointly upon the final number record and their contributions are not known individually at each instant. Hence substantial corrections are not feasible. An exhaustive enumeration of factors does not make foreign trade data, when they have the properties listed by these factors, suitable for a fine-structured theory of international trade using such concepts as terms of trade or for monthly comparisons with exchange rates, interest rates, etc.

The most striking example is offered by gold, a clearly defined, easily recognizable commodity ,of high value relative to quantity, costs of shipment and usually accorded exceptional care in handling. Yet statistics of gold movement, pairwise compared, reveal disastrous differences whether 1900, 1907 or 1928, 1935 are taken as samples. This holds true for monthly as well as cumulative yearly data.

There appears to be little hope ever to reconstruct series for the past on gold useful for most purposes of economic theory. Present information can only be correlated with the discussion of broad tendencies in economic life. The same is true, *a fortiori*, for the far more involved statistics of other commodities and the large categories into which they are put. These sober facts should lead to the establishment of a program for future data on foreign trade with known errors, perhaps on a sampling base.

On Respondent-Nonrespondent Differences Observed in the Pittsburgh Morbidity Surveys. DANIEL G. HORVITZ, *North Carolina State College.*

In morbidity surveys in particular, and other surveys as well, it is not uncommon to find a field procedure in use which permits the interviewer to obtain information required on *all* members of each selected household from any responsible member who happens to be available at the time of the call. The efficiency of this procedure is open to question whenever the information supplied by the respondents on the other individuals (nonrespondents) is inaccurate.

In the absence of an opportunity to measure the response bias directly, data collected on illnesses experienced in the month prior to interview in two surveys (carried out a year apart with a probability sample of approximately 3000 households located in the Arsenal Health District of Pittsburgh) were subjected to internal study in an attempt to characterize the rather large differences in the rates observed for respondents and nonrespondents. The data were examined first to determine the relationship, if any, of these differences to age, sex and size of household. Almost in general the data yielded no indication of a relationship between the observed respondent-nonrespondent differences and these variables whether for all illness or specific categories of illness. There was slight evidence of a greater respondent-nonrespondent difference (i) for males than females for noninfectious conditions and (ii) for those in the 15–34 year age group versus older age groups for all types of illness. This latter observation was accounted for by a strong relationship between respondent-nonrespondent differences and age for female conditions associated with pregnancy and childbirth. For these conditions the respondent-nonrespondent differences can be considered actual rather than due to any reporting bias. The respondent-nonrespondent prevalence rate differences observed for the specific chronic conditions of arthritis and heart disease also were found to be unrelated to age, sex and size of household.

Examination of the respondent-nonrespondent differences for illnesses still in progress on the date of interview versus those which had already ended indicated that a reasonable portion of the total difference may be an actual fact. This observation together with the consistency of the differences with age, sex and size of household requires further verification from morbidity studies conducted elsewhere. They also require (pending substantiation) investigation of the feasibility of a scheme for morbidity surveys whereby in a relatively small random segment of the sample households the respondent is selected at random from those eligible. The respondent-nonrespondent differences in these households should provide an estimate of the reporting bias portion of the differences observed in the remaining households. This suggestion presumes that the actual and bias portions of the differences do not vary in a compensating fashion to yield total differences unrelated to age, sex and size of household

Questionnaire Design and Related Methodological Problems in the Canadian Sickness Survey. ROBERT KOHN, *Dominion Bureau of Statistics, Ottawa.*

Problems of design are somewhat different for questionnaires and schedules, although many requirements apply to both. The basic document in the Canadian Sickness Survey was an interview schedule. Principles of design are discussed in three main categories: 1. contents, 2. wording, and 3. layout and format. The solution for particular problems will always depend on the type of informant and enumerator to whom the questions are directed.

The contents are dictated by the objective of the Survey and the tabulation program. They should be kept to the minimum required. If after the beginning of the survey the need arises to add supplementary questions, these should not be allowed to jeopardize the original ones and enumerators should be fully informed regarding their purpose and meaning. The wording must be such as to elicit the same answer in the same situation from a variety of people. It must be adapted to those of the prospective informants and enumerators who know and understand least about the survey. It must be easily understandable and yet concise. Instructions must also be drafted with great care and some machinery must be provided for prompt and uniform explanation and supplementation of the instructions if required during the survey. Layout and format must be such as to facilitate interviewing, recording, checking, editing, and coding. Pretesting of schedules will reveal shortcomings before the actual survey gets under way. Qualification and training of enumerators are important and so are methods to solicit and maintain the co-operation of the informants.

The Future of Railroad Shares. PIERRE R. BRETEY, *Heyden Stone & Co.*

During the past decade the rails have suffered from three major factors: (1) Delays by regulatory authorities in adjusting freight and passenger rates to higher costs, particularly following sizable wage increases. There have been many such delays since World War II. However, the "adjustment lag period" has consistently been reduced over recent years, and may be almost entirely eliminated through passage of legislation. (2) The inflexibility of a large percentage of railroad costs, with the result that historically earnings have declined sharply when gross revenues contracted. Notwithstanding this inflexibility, a number of railroads had achieved a flexible cost position before the roll-over adjustment of 1949 and demonstrated better expense controls than exhibited by many industries. (3) A lowered institutional appraisal resulting from the collapse in railroad earnings and in railroad equities both in 1929–1932, and again in 1938, collapses not duplicated since those two periods. Today most investment railroad equities are selling at from only four to five times estimated 1953 earnings, and railroad managements agree that individual properties are in the best physical condition in their history and that operating efficiency of most railroads has been restored following the rapid increase in costs immediately following

the end of World War II. They are now entitled to be regarded as akin to many other industries sensitively dependent upon the strengthened ability of individual managements in adjusting to changing conditions. I confidently expect that budgetary controls, maintenance cutbacks and more efficient operations resulting from cumulative benefits of some $9 billion of capital expenditures made since V-J Day, will maintain 1954 earnings of the Class I railroads at between $600 and $700 million. We are confident that after the lapse of fifteen years or more, institutional investors will once again reenter the railroad market. If this demand does make its appearance, it may well serve to spark another major market leg in the long road toward ultimate restoration of railroad credit to its once former high estate

The Changing Stock Market. SIDNEY LURIE, *Paine, Webber, Jackson & Curtis.*

There are a number of inter-related factors which have an important bearing on the outlook for our economy and, in turn, the stock market: (1) The population growth already experienced and still in prospect is opening up new markets for American industry—for the gains are in the ranks of the consumer rather than producer. (2) The redistribution of our national income, with the greatest increase being in that segment which proportionately spends the most, injects a new stabilising force. (3) America's inventive genius in developing new products which open up new markets has spelled profits for the investor as well as manufacturer. While these considerations give rise to an era which historians may term "The Fabulous Fifties," two other factors have created an enigma for the security buyer: (A) Although common stocks today are more respectable than ever before in Wall Street history, the public has lost sight of the fact that the rewards always go to the risk-taker. The security buying public has forgotten there is nothing illegal, immoral or fattening to speculation—that there is little basic difference between intelligent speculation and intelligent investment. (B) The big financial news of the postwar years is the creation of a new class of security buyer: the indirect investor. The growth of pension funds and investment trusts has resulted in an "institutionalized" stock market. The foregoing introduces a new element of stability to the price level—but it also means that: (1) The dominant buying interest is confined to a relatively small portion of the available issues (2) A disparity in values has been created which eventually must be corrected. If the interest of the security buying public can be revived everyone will benefit—industry as well as the investor. The broader the demand, the more likely all stocks will reflect their basic values. An active stock market is synonymous with a vigorous economy.

Trends in Monetary Structures in 1954. M. DUTTON MOREHOUSE, *Brown Bros., Harriman & Co.*

In the nonbank long-term market, the supply of funds could well exceed the private demand by some $4 billion. In the short-term bank market the supply of funds will be governed by Federal Reserve policy, but there should be a negative net private demand of some $2½ billion. The Treasury should be able, therefore, to achieve measurable progress in extending the maturity of the national debt without running the risk of upsetting the market. The Federal Government might show a surplus of receipts over expenditures in the first half of 1954 in the magnitude of $7 billion and a deficit of $10 billion in the second half, or a net deficit for the calendar year of $3 billion. Administration efforts to reduce expenses might well result in a smaller deficit.

The decline in private demand for bank credit must be replaced by Government credit to avoid the deflationary pressure of a decrease in money supply. It is evident, therefore, that the bank market not only can but should absorb an amount of Government securities equivalent to the estimated deficit for the calendar year 1954. For the first time in many years the Treasury will find a market receptive to the offering of long bonds, and yet monetary policy appears to dictate the placing of a substantial amount of additional Treasury securities within the banking system. A ready market should be found within the banking system for intermediate maturities at yields equivalent to or somewhat lower than those prevailing today. Throughout the year the money market will be one in which a surplus supply of private funds seeking investment will be pushing towards lower yields. Frequent Treasury offerings will be designed to absorb these surplus funds as they accrue and thus prevent yields from changing substantially. Nevertheless, higher prices and lower yields are expected to prevail a year from now.

Monopoly, Bigness, and Progress. G. WARREN NUTTER, *Yale University.*

It has been often and widely argued that monopoly and big business are not only an inevitable consequence but also a necessary stimulant of economic progress. There is an element of truth in this proposition, but it emerges only after the argument has been severely qualified. In the first place, active competition can and does persist in industries characterized by continual growth and innovation Moreover, for the economy as a whole competition has scarcely diminished in vigor or extent in the face of the rapid and continual growth over the last half century; if anything, it has increased. In the second place, friction, in the form of delayed reactions to the lure of high profits, is the condition needed to induce innovations; and there is enough friction in most competitive industries to call forth

a constant stream of new products and technological improvements. Some of the outstanding new products of recent decades, like the automobile and radio, were introduced into competitive surroundings. Monopoly makes possible, though it does not ensure, extraordinary friction; and therefore it increases the likelihood that risky and costly innovations will be made. Large size combined with monopoly, a combination that is not inevitable, encourages the most risky and costly innovations. We now know very little quantitatively about the respective roles of competition, monopoly, and big business as promoters of economic progress. If discussion is to lead to intelligent changes in public policy, we must improve our understanding of these roles together with our understanding of how organizational forms affect other important values of our democratic and libertarian tradition.

Validation of Morbidity Survey Data by Comparison with Medical Records. NEDRA B. BELLOC, *California State Department of Public Health.*

A household sample survey of morbidity in the general population was conducted in San Jose by the California Department of Public Health in the spring of 1952. The survey was designed to test several methods of measuring morbidity and to study the completeness and accuracy of the information obtained. For validation purposes, medical records were collected from a number of sources. With respect to hospitalized illness, it was possible to study the net error in reporting. Admission rates based on household survey reports did not differ significantly from those based on hospital records for the same population. Similarly, days of hospitalization per person per year and average length of stay per period of hospitalization were measured accurately from household survey data. Distributions of admissions from household survey reports of hospitalization by month of admission, length of stay, and diagnosis were similar to the distributions obtained from hospital records. Whether or not surgery was performed was reported accurately in the household survey, but the description of the surgical procedure was not as precise as that obtained from hospital records. This study shows that reports of hospitalization obtained in a household sample survey are sufficiently accurate to be used for many purposes in lieu of hospital record data.

Hospital Morbidity Reporting—Experiences and Findings of a Pilot Project. I. Methodological Experiences. CARL L. ERHARDT, *New York City Department of Health.*

The Hospital Morbidity Reporting Project had two major aims: to test a method for collecting and tabulating data regarding patients discharged from hospitals in the city, and to analyze a meaningful body of the data collected. Both general and special hospitals, operated under municipal and other auspices, participated. Data covering the complete hospital system operated by the city are available for analysis. Reports submitted for each patient discharged during the project period included the usual demographic descriptions plus medical diagnoses, data on operative interventions and length of hospital stay. The facts, including diagnostic and operative data, were reported in code or in self-coding form, eliminating the usual expensive central processing step. The reported diagnostic code (Standard Nomenclature of Diseases and Operations) was mechanically translated into a statistical grouping of diagnoses (International Statistical Classification of Diseases, Injuries and Causes of Death) for analysis. The punching of digits of the Standard Nomenclature code was limited to those necessary for ISC classification. A condensation of the ISC was also designed for summary tabulation of hospital diagnoses significant in frequency in the New York City area. The methodological experiences and analytic procedures of this pilot project provide necessary background on which to base plans for routine collection of hospital morbidity data in New York City.

Hospital Morbidity Reporting—Experiences and Findings of a Pilot Project. II. Significance of Findings. MARTA FRAENKEL. *New York City Department of Hospitals.*

Medical diagnoses together with demographic data on hospital patients can contribute to morbidity statistics needed by public health administration. In conjunction with information on duration and outcome of hospital care, these morbidity data have significance for hospital administration and medical care planning. To date, most hospital statistics cover all patients disregarding their diagnoses. The analyses of the Project data showed that the composition of the patient load by age and diagnosis of 14 municipal general hospitals differed substantially. These differences explain many deviations in the data on "average length of stay," "net death rate," etc. if computed in the traditional unspecific way. It is shown how disease-specific data would permit more meaningful evaluation of hospital services and provide more useful bases for estimating personnel needs. Project data supplied disease-specific information on chronic illness. Many chronic disease patients stay only shortly in general hospitals. Short-stay hospitalizations—for special diagnostic, therapeutic or rehabilitative measures or for care during acute exacerbations—are elements of a long-term medical regime of many of these patients. Community-wide organization of integrated medical programs for chronic illness requires current disease-specific information on hospital care to these patients.

Statistical Determination of Tolerances in Rocket Development. EDWIN L. CROW, *U. S. Naval Ordnance Test Station.*

The paper considers the problem of determining the tolerances necessary for the components of a rocket in order that it have a prescribed dispersion. The establishment of the relation between departure from nominal construction and the resulting deviation of the rocket from the ideal trajectory is the main concern of this study, but two other requirements and their utilization are also discussed briefly: the probability distribution of the departures under manufacture within any given tolerances, and the relation between any particular tolerance and the cost of attaining it. To determine with satisfactory precision the relation between the deviation in trajectory and departure in construction of a particular rocket, an unusually large departure was purposely introduced into 50 rockets, which were then fired along with a control group. Firing was repeated with the departure doubled, and several types of departure were considered. In general the departure affected not only the mean coordinate of the trajectory but also the variances. A maximum likelihood estimate of the magnitude of the deviation was not practically attainable. An unbiased estimate of the mean square deviation was obtained by the method of moments, and its variance derived and investigated for various cases.

The Up-and-Down Method with Small Samples. J. L. HODGES, JR., *University of California (Berkeley)*

The main content of this paper is a report on computations made to determine the actual performance of some estimates for the mean dosage parameter μ, based on up-and-down series of length 10 or less. It is found that the Dixon-Mood formula for the asymptotic variance is reasonably reliable even in samples as small as 5 to 10. Thus, restriction (b) is not necessary. As a consequence of this fact, the design of up-and-down experiments may be altered, so that several independent series are run simultaneously, without serious loss of accuracy. In this way, restriction (a) can be considerably reduced. Finally, the possibility of conducting an experiment with several short series run in parallel introduces a flexibility into the design, which permits us to take advantage of special features of some experimental situations with a still further increase in efficiency.

Use of Statistics in Engineering in France. ANDRÉ G. LAURENT, *University of Chicago.*

In France, statistical applications in engineering are chiefly concerned with quality control with stress on: a) teaching, evidenced by creation of a training center in statistical methods for engineers at the University of Paris; b) publications, evidenced by books by Dumas, Laurent, Mothes, and a periodical, *Revue de Statistique Appliquée*, c) applications to industry, evidenced by applications in mining, mechanical, metallurgical, and textile industries. Difficulties arise from lack of financial support, poor reputation of statistics among industrialists, and the gap between theory and practice. Further developments may be expected along several directions: a) widening the fields of application; b) shifts of emphasis to descriptive and praxeological aspects as well as upon inference; c) more widespread use of available but little applied methods.

An Example of a Fractional Replication in a Bearing Abrasive Wear Test. F. R. DEL PRIORE and WILLIAM J. KOMMERS.

The purpose of this study was to develop an accelerated test for determining the relative wear resistances of plastic bearing materials. First it was necessary to find the effect of the test conditions on bearing wear. Physical considerations and limitations played the usual important role in the testing schedule. Six variables, each at two levels were investigated as a quarter-fractional. Three bearing materials were tested as split plots. The details in setting up the test plan are given as well as the analysis of the data. Following the analysis of the first quarter-replicate, it was considered that adequate results were obtained to obviate the testing of additional quarters. A special bench testing machine was designed and built on the basis of this work. Fixed test conditions were selected that produced the maximum differentiation between the wear resistances of the materials in a given period of time. The machine makes it possible to evaluate bearing materials in a matter of hours instead of the former month-long simulated service test procedures.

Gasoline Mileage in Winter Day-to-Day Use. DAVID FRAZIER and RAY DECKER.

Research executives of a petroleum company suggested that mileages of motor gasolines in northern winter "home to work" driving might differ importantly because of differences in starting losses, or in warm-up and misfiring characteristics. Such differences, if they really existed and were of important size, might suggest how the company could gain a competitive advantage. Therefore mileages of 15 gasolines, 7 commercial and 8 experimental, were compared in a complex statistical experiment on 15 employee owned and driven cars during the winter driving season. Each car in the course of the experiment ran 7 of the 15 gasolines each for one week, the gasolines being assigned to cars according to a statistical

plan called a "15 by 7 Youden square." All cars were driven normally, and gasoline consumption recorded for each gasoline on each car was the difference between gasoline added and gasoline drained at the end of the week.

When the experiment was complete statistical analysis, which eliminated all inherent "car to car" differences in gasoline mileage and much of the "week to week" variation within each car, revealed no significant differences among the gasolines. Further, the 8 experimental gasolines were so compounded that they constituted a 2^3 factorial which, when analyzed, showed with improved power that variation of vapor pressure, olefin content, or polyolefin content within reasonable limits had no significant effects on winter gasoline mileage. Inspection of results showed that real differences among the gasolines large enough to be important commercially would almost certainly have been reported as significant in this experiment, and therefore it was concluded that gasoline composition within reasonable limits has no commercially important effect on winter gasoline mileage.

Operations Research as a Science. JOHN B. LATHROP, *Arthur D. Little, Inc., Cambridge, Mass.*

During the last few years, the term "operations research" has been applied to a growing body of investigative activity, directed at problems of decision and management in many areas of business, industry, and government. It has been described as the application of the scientific method in new areas of decision, with the broadening of the range of such application of such magnitude as to require the appellation of a new branch of science. Operations research is certainly closely related to other fields of science and engineering, such as systems engineering, industrial engineering, market research, etc., and is often concerned with problems of interest to statisticians. These relationships can be illustrated by pointing out some of the general characteristics typical of operations research, and the role played by statistics in some examples.

As in science, the primary objective of operations research is to understand, not to act. When an operation is understood, the action required to improve the process is often fairly evident. The scientific method, by a combination of quantitative hypothesis, observation, and controlled experiment provides this understanding by means of relatively simple models to describe complex situations. Rather than reasoning from the facts to the mechanism, as in some forms of statistical analysis, the operations research scientist sets up assumed models in order to deduce phenomena to check against observed facts. An excellent and well known example of this method was Newton's explanation of the apparently unrelated phenomena of planetary motion and objects falling on the earth, by the simple unifying concept of gravity. In a study to reveal the causes of corrosion in paper mill equipment, the number of possible factors—chemical, physical and metallurgical—was very large, as was the amount of observed data from what amounted to an uncontrolled experiment. Although variance analysis was useful in determining the relative importance of various classifications of corrosion, statistical analysis alone would reveal no causes. It was necessary to develop a model of corrosion from first principles, using the physical sciences, in order to obtain a useful understanding of the situation.

Other models have revealed the relationships between sales promotion effort and sales, between production schedules and production costs, between warehouse stocking procedures and service to customers, between salesmen's compensation plans and profits, and so on. In some, statistical methods played an important part, in others no part at all. Information theory, queuing theory, servomechanism theory, symbolic logic or simple punched card reproductions of an operation have all provided the basis for operations research models. To summarize, operations research is a branch of scientific research concerned more with reaching an understanding of an operation than with the relations among the numbers describing the operation.

Effect of Current Operating Experience on the Realization of Investment Plans. JEAN BRONFENBRENNER CROCKETT.

Comparison of the planned with the actual fixed capital outlays of manufacturing firms, as reported to the Department of Commerce and the Securities and Exchange Commission, has indicated that individual firms' deviations from annual investment programs are subject to large variance. The hypothesis is tested that some significant part of the variance can be explained by concurrent fluctuations in the sales, profits, and liquidity of the investing firms. These variables may influence investment decisions either by affecting the expected return from proposed investment (in the first two cases) or by affecting the ability to finance desired expenditures (in the third case). The data studied cover about 400 firms in the mild recession year of 1949. Information on actual and anticipated investment and sales was supplemented by income and balance sheet data relating to profits and liquidity.

The extent and nature of the effect on investment decisions of the explanatory variables may be expected to depend on whether the proposed investment is primarily for expansion or primarily for modernization. Direct information was not available, but a rough attempt was made to separate firms which might be, from those which probably were not, interested in expansionary investment. For the

first group, deviations from investment programs were found to be correlated positively and very significantly with deviations of actual from anticipated sales, while for the second group the effect of sales deviations was not significant.

Changes in profit rates, as compared with the previous year, were also found to have a significant effect on the investment deviations of "expansionary" firms, but this was less marked than the influence of sales deviations and impossible to separate from it because of high correlation between the two explanatory variables. For the "non-expansionary" group, the effect of profit movements was not significant, and (except for cases of very large profit declines) the relationship appeared to be negative, possibly indicating a tendency for modernization expenditures to vary inversely with profit movements arising from changes in labor and materials costs. The effect of changes in liquidity on investment deviations is difficult to ascertain. However, in those cases where the observed liquidity movement is clearly something more than a *result* of the deviation from investment program, it appears to be a significant factor in *motivating* the deviation.

Expectations, Plans and Capital Expenditures: A Synthesis of Ex Post and Ex Ante Data. ROBERT EISNER, *Northwestern University.*

Ex ante and *ex post* data relating to sales and capacity changes, profits, capital expenditures and other variables have been derived from replies in McGraw-Hill questionnaire surveys and from independently obtained corporate balance sheet and income statement information. Analysis is focussed on the determinants of investment and, in particular, on evidence of operation of the acceleration principle. An "acceleration component" of investment is isolated and observed. Its absolute magnitude is found to be positively related to both current and lagged actual changes in sales and to expected changes in sales. And a larger "acceleration component" is associated with a larger total of investment. This evidence seems to fit well assumptions of recent economic theory of Hicks and Harrod which suggests a great significance for even a multilagged and quantitatively limited role of the acceleration principle. Investment is also found to be positively related to the rate of profit.

Expectations of sales increases are related to very great or long-term growth, but current sales changes are generally related negatively or not at all to expectations of immediately future changes. The directions of expected changes, but not their magnitude, can be matched to corresponding actual changes in sales. Discrepancies between planned and realized investment are found to be a decreasing function of size of firm, with very large firms appearing slower to adjust in a period of rising expenditures. Sales changes different from those expected, particularly in the case of firms expecting increases, are found to be a factor in the difference between investment plans and realizations.

Teaching Statistics to Executives. W. ALLEN WALLIS.

My discussion is based on experience in teaching statistics in the Executive Program at the School of Business at the University of Chicago. The Program's students are experienced executives. The objectives of the course are to teach: (1) avoidance of common logical mistakes in reasoning from data; (2) techniques for organizing numerical data into comprehensible form; (3) appreciation of the problems of allowing for chance and for unanticipated effects in designing investigations and in drawing conclusions from them; (4) a few rough-and-ready methods of analysis; (5) an appreciation of the potential uses of sampling; (6) recognition of a statistical problem when encountered; and (7) respect for the intellectual discipline underlying modern statistics.

Although the course operates mainly through lectures, the atmosphere is informal; class members raise questions and make many contributions, and time is left for post-lecture discussion. Numerous realistic examples are used, not only from business but from other fields. While we try to bring out the general principle illustrated by an example, a systematic exposition of statistical principles would have fairly low priority. What the students want, and what we are trying to give them, is something which will enable them to understand and generalize the experiences they have had so they may better cope with the broader experiences they expect to have. An important teaching device is the term paper. The students submit written descriptions of a statistical problem (preferably connected with their own business) and of the methods they propose to use for attacking it. This is returned to them with comments and suggestions, and must be approved by about the middle of the term. The final papers, submitted at the end of the term, are read carefully, and many students receive letters from the instructor commenting on them in detail. A statistics graduate student attends the lectures and writes detailed notes for distribution not later than the next class meeting. Regular textbooks are also used, and some exercises and quizzes are given.

We have come to feel that the kind of course developed for the Executive Program is more suitable on the campus than the kind of introductory course traditionally given. Consequently the courses in the Business School and in various other departments are being unified along the lines developed in the Executive Program.

Intercensal Needs for Small Area Data—By a Local Planning Agency. HARLIN G. LOOMER, *Philadelphia City Planning Commission.*

The statistical information that is needed by a local planning agency during the intercensal periods may be divided into two classes: that which is available or can be developed from local sources, and that which requires the special resources and authorities of a federal agency for collection and development. The former should be fully exploited before federal agencies are called upon for assistance. In the present situation of curtailed budgets, federal agencies should not undertake a multiplicity of new field survey projects, even on a sample basis, but should concentrate on fully exploiting and coordinating material that already is being collected by several such agencies. If all of the usual complete enumerations of the Census and all of the continuous records that are now being maintained by such agencies as B.L.S., Federal Reserve Bank, Old Age Assistance, etc. were properly brought together and integrated, there would be sufficient indices to indicate the trends from one decade to the next. Federal-local cooperative research should be further developed, possibly through the expansion of the facilities of regional or local branch offices of the federal agencies. Trained and experienced personnel from these offices could render valuable service in planning and directing local research. As a permanent continuing agency, these offices could bridge the gaps between intermittent or one-time surveys by local agencies and coordinate all into a comprehensive whole for the area. The authority and reputation of a federal agency (such as the Census) is needed to assure the success of some types of studies. Costs should be shared by participating agencies. Uniform sample surveys are of little value to local planners. Either the sample is too small to indicate purely local trends or the procedure is too standardized to meet specific local needs. A highly flexible cooperative arrangement would best serve the needs of local planning agencies.

Business Uses of Intercensal Data. W. R. SIMMONS.

Excepting local establishments, it is safe to venture that there are few business firms which do not make many vital planning and operational decisions based directly or indirectly on the findings of the Census. Generally taken for granted, like the mail and telephone service, Census reports become conspicuous only when and if they are absent. Not the least of the considerable American economic progress has been due to more intelligent planning based on sound factual information, to which Census reports are indispensable. In establishing plant and store locations, in developing local, regional, and national sales potentials, in defining sales territories, in analyzing problems of distribution, in evaluating sales performance and in deciding where, when and how to improve future performance, this type of information becomes crucial. Any substantial reduction in this service, in the commendable interest of economy, must run the risk of impairing the efficiency of our entire economy.

Meeting the Needs for Small Area Intercensal Data. CONRAD TAEUBER.

The major method of meeting needs for intercensal data for small areas from Federal sources is through Special Censuses conducted at the request and expense of local areas. Nearly 200 such Censuses have already been conducted since 1950, primarily for the purpose of establishing base figures for the allotment of tax funds. Although considerable interest in a quinquennial census of population about 1955 has been expressed, the prospects for it do not seem bright because of the costs. Estimates of population change since the last census and projections into the future are prepared locally by many cities and counties. The Census Bureau has issued a report suggesting methods of estimating current populations of small areas. School censuses are well established, but generally are not conducted in a manner to provide general purpose statistics. They represent a major resource for local area data which needs further development. There is need also for further research into the utility of the wide variety of local sources which could be used as indicators of population change in individual localities.

Industrial Production Index. CLAYTON GEHMAN, LORMAN TRUEBLOOD, ARTHUR L. BROIDA, M. H. SCHWARTZ, MILTON MOSS, PETER M. CODY, *Federal Reserve Board.*

The Board of Governors of the Federal Reserve System on December 18 released the basic postwar revision of the index of industrial production. The more obvious major features of the revision are an up-dating of the comparison base period from 1935–39 to 1947–49 and a change to the Standard Industrial Classification from the prewar Census classification system. More substantive features relate to the up-dating of the weight year from 1937 to 1947, enlargement of the number of monthly series from 100 to 175, substantial improvement of many old series, the development of independent annual indexes to which the monthly indexes are adjusted, and new seasonal factors for all major groups.

The revision so far applies mainly to the period from 1947 to date. For the earlier period the Board has adjusted its old manufacturing output indexes for the years 1939–1947 to the comprehensive benchmark changes developed by a joint Census-Federal Reserve project, and has linked the old to the new in January 1939 for the period back to 1919. Similar benchmarks and links were made for minerals

These interim revisions have been done at the level of the major divisions of the index in order to facilitate comparisons with the more recent period. In a general way, changes in industrial activity since 1947 are shown to be similar by both the new and the old total indexes. In the first half of this year, both indexes indicate that activity was at a record level for the postwar period, about one-eighth above a year ago. Both show that since midyear, output has been reduced fairly generally. For October, the new index was about 4 per cent below the highs which it established in May and July; the old index was down about 5 per cent from its peak reached in March. Both indexes show that industrial production in October was at about the same level as a year earlier.

Demonstration Teaching of Statistics. R. E. WAGENHALS.

You can interest people in the use of statistical quality control without the use of mathematics by dramatic factual demonstrations. With the use of a series of demonstration devices, people can be shown how ineffective 100% inspection and 10% sampling may be as compared to scientific or statistical sampling. The repeated winner of a modified game of dice demonstrates he is not lucky but is just using control chart principles. A small pin ball machine demonstrates how a control chart and the normal distribution is related. The advantage of having quality controlled products is demonstrated by five different colored groups of twenty tubes to show how assembly tolerances can be reduced. These devices are demonstrated and the source of information regarding their construction will be available.

Use of Experiments in Engineering Statistics. IRVING W. BURR, *Purdue University.*

The paper presents convenient populations for measurements and binomial and Poisson distributions. Then experiments to illustrate the following are discussed: sample frequency distribution, control charts, significance tests, estimation, analysis of variance, statistics of combinations, acceptance sampling, sequential analysis, linear correlation, and curve-fitting. A simple method of simulating a normal bivariate population is presented.

Using the Experimental Approach in the Teaching of Statistics. EDWIN G. OLDS, *Carnegie Institute of Technology.*

After a brief discussion of types of experiments which might be used, the paper discusses the value of the experimental approach in connection with lectures by the instructor and laboratory work by the student. The importance of using adequate time for orientation is emphasized. The paper closes with a proposal that the opinions stated be put to test by means of a designed experiment.

Problems in Experimenting with the Application of Statistical Techniques in Auditing. JOHN NETER, *Syracuse University.*

Sampling is used extensively in auditing. Before any attempts to experiment with the use of statistical techniques in auditing can be made, one must formulate the purpose of sampling in auditing in such terms that statistical models can be constructed either for the problem as a whole or for definite parts of it. Once this has been done, the relative merits of various statistical models can then be investigated. One basic approach to the purpose of sampling in auditing would be to ascertain whether the figures in the financial statements are reasonably close to the figures which would be obtained if a complete examination of all transactions and records, including any necessary corrections, were made. Another approach would place emphasis upon an examination of the basic company accounting policies and of the effectiveness of the internal control process in assuring the accurate recording of transactions in accordance with these policies. Still another approach would encompass both of these points of view Once the purpose of sampling in auditing has been suitably formulated, problems such as the determination of the alternative decisions from which a choice is to be made, the criteria according to which the appropriate decision is to be chosen, the relevant statistical model for the population under study, and the most appropriate sampling procedure can be studied and appropriate experiments conducted to determine the suitability of various statistical approaches.

A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances. ROBERT BECHHOFER, *Cornell University.*

A Two-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with a Common Unknown Variance. CHARLES W. DUNNETT, *Lederle Laboratories.*

A Sequential Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances. MILTON SOBEL, *Cornell University.*

A multiple decision approach is considered to the problem of partially or completely ranking a given set of population means. The particular goal of finding the largest of k population means is treated in detail. Let $\delta_{k,k-1}$ denote the true, unknown difference between the largest and second largest populations means. It is assumed that the experimenter can specify before experimentation starts: (1) The smallest

value, say δ^*, of $\delta_{k,k-1}$ that he is interested in detecting and (2) The smallest acceptable value, say P, of the probability of achieving this goal when $\delta_{k,k-1} \geq \delta^*$. Three multiple decision procedures to accomplish this goal are given: (1) A single-sample procedure when the common variance is known. (2) A two-sample procedure when the common variance is unknown. (3) A sequential procedure when the common variance is known The observations X_{ij} are normally and independently distributed chance variables $N(X_{ij}|\mu_i, \sigma^2)$, $(i=1, 2, \cdots, k; j=1, 2, \cdots \text{ad. inf.})$. We assume that the μ_i are unknown. The ranked μ_i are denoted by $\mu_{[1]} \leqq \mu_{[2]} \leqq \cdots \leqq \mu_{[k]}$. It is not known which population is associated with $\mu_{[i]} (i=1, 2, \cdots, k)$. We define $\delta_{ij} = \mu_{[i]} - \mu_{[j]} (i, j = 1, 2, \cdots, k)$.

Single-Sample Procedure. (1) Enter the appropriate table with k and the specified P, and obtain a constant $h = h(k, P)$. (2) Set h equal to $\sqrt{N}\lambda$ where $\lambda = \delta^*/\sigma$ and σ is the *known* population standard deviation. Solve this equation for N. (3) Take N observations from each population where N is the smallest integer greater than or equal to the solution obtained in (2). (4) Calculate the k sample sums. (5) Select the population which yielded the largest sample sum as the population having the largest population mean.

Two-Sample Procedure. (1) Take a first sample of N_0 observations from each of the k populations. Any integer N_0 will satisfy the requirements of the problem. (2) Calculate

$$s_0^2 = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{N_0} \left(X_{ij} - \frac{1}{N_0} \sum_{j=1}^{N_0} X_{ij} \right)^2$$

which is an unbiased estimate of σ^2 having

$$n = k(N_0 - 1)$$

degrees of freedom. (3) Enter the appropriate table with $n = k(N_0 - 1)$ and the specified P, and obtain a constant $h = h(n, k, P)$. (4) Take a second sample of $N - N_0$ observations from each of the k populations where N is the smallest integer equal to or greater than the larger of N_0 and $2s_0^2(h/\delta^*)^2$. (5) Calculate the k over-all sample sums $Y_{iN} = \sum_{j=1}^{N} X_{ij}$ $(i=1, 2, \cdots, k)$. (6) Select the population which yielded the largest over-all sample sum as the population having the largest population mean.

Sequential Procedure. Let the ranked sums based on m observations be denoted by

$$Y_{[1]m} < Y_{[2]m} < \cdots < Y_{[k]m}$$

and let

$$D_{im} = Y_{[k]m} - Y_{[i]m} \qquad (i = 1, 2, \cdots, k-1)$$

"At the mth stage of experimentation take an observation from each of the k populations and compute

$$W_m = e^{-\delta^* D_{1m}/\sigma^2} + e^{-\delta^* D_{2m}/\sigma^2} + \cdots + e^{-\delta^* D_{k-1,m}/\sigma^2}.$$

(a) If $W_m \leqq (1-P)/P$, stop experimentation and choose the population which yielded the largest sum, $Y_{[k]m}$, as the one having the largest population mean. (b) If $W_m > [(1-P)/P]$, take another observation from each of the k populations and compute W_{m+1}. Continue in this manner until the rule calls for stopping."

An Optimum Slippage Test for the Variance of K Normal Populations. DONALD TRUAX, *Stanford University.*

A solution is given to the problem of deciding if several normal populations with unknown means and variances have equal variances, or if not, which population has the largest variance. Under some mild restrictions on the class of possible decision procedures this solution is shown to be optimum in the sense that it maximizes the probability of making the correct decision when all the variances are equal except one which is larger than the rest. The optimum procedure is given as follows: select population Π_M if $s_M^2 / \sum_{i=1}^{k} s_i^2 > L_\alpha$ and decide that all the variances are equal if $s_M^2 / \sum_{i=1}^{k} s_i^2 \leq L_\alpha$. We denote by s_i^2 the unbiased estimate of the variance from the ith population, and M represents the index of the population having the largest sample variance. The constant L_α is determined by the restriction that if all the variances are equal then that decision should be made with probability $1-\alpha$.

The Use of Statistical Techniques in the Aging of Accounts Receivable. R. M. CYERT, *Carnegie Institute of Technology.*

The aging of accounts receivable in most enterprises is done by sampling a portion of the accounts receivable. On the basis of the aging an allowance for uncollectibles is estimated. The study described in this paper had two primary objectives: (1) To analyze the sample design being used for selecting accounts to be aged with a view toward improving the method of selection and determining the optimum sample size. (2) To examine the plan for selecting accounts already aged to be test-checked by the auditor

with a view toward installing an appropriate acceptance sampling scheme. The procedure followed was to examine a past sample drawn from a large metropolitan department store. The sample was first analyzed for randomness by the use of Kendall's coefficient of rank correlation. There was no significant correlation found between the alphabetical rank of the accounts and the rank of the account age. It was decided, therefore, that it would be meaningful to compute the precision with which the past sample estimated the allowance for bad debts. The precision as computed from the sample of 15,000 accounts was deemed too tight by the public accountants working on the study. New precision and reliability requirements were specified and a new sample size was computed on the assumption of an unrestricted random sample design. The required sample size was reduced from 15,000 to 1,700. The sample design specifies that the accounts will be drawn systematically and the sampling error will be computed by the Tukey plan.

Previous practice in test-checking the agings done by store personnel had been to select 10% of the accounts aged. The public accountant then decided whether or not to accept the total accounts aged on the basis of the test-check. The process of test-checking was one to which sequential analysis was applicable. After the accountant had defined a good and a bad lot and specified the risks to be taken, a sequential sampling plan was determined. The plan was truncated at three times the average sample number for lots with a per cent defective equal to the slope of the decision lines. The maximum number that would need to be sampled, therefore, is 376 and the average sample number is 126. In the authors' view, this study establishes the possibility of adapting statistical methods to accounting and auditing problems involving a high volume of homogeneous data. It is important to note that the subject of this study and, therefore, its conclusions do not relate to an area of audit in which results are applied directly to the decision of giving or withholding an opinion on financial statements.

New Experimental Designs for Paired Observations. W. S. CONNOR, *National Bureau of Standards.*

There are many experimental situations in which only two objects can be observed at a time under homogeneous conditions, as for example in a recent experiment at the National Bureau of Standards which involved the comparison of several types of spark plugs in two cylinder gasoline engines. In these cases experimental designs with two plots per block are needed. If all possible pairs of objects are observed, the statistical analysis of the observations is simple. However, if the number of pairs is large, this procedure may be expensive and may provide more precise results than are needed. Thus it often is desirable to observe only part of the pairs. The choice of this subset of pairs must be made with care if the statistical analysis is to remain simple. Such a subset is the "two-group arrangement," which was suggested by the traditional experimental procedure of comparing several new objects with one or more standards. The two-group arrangement consists of dividing v objects or treatments into two groups of m and n treatments ($v = m + n$), and of pairing every treatment from one group with every treatment from the other group. No other pairs are formed. The analysis appropriate to this arrangement is described in detail.

Some Aspects of Sequential Experimentation. S. MUNRO.

Out of practical necessity a sizable amount of industrial and military development is being conducted sequentially, even to the extreme of a sequence of samples of one. There is, however, aside from a handful of exploratory and superficial papers, no body of mathematical theory which serves as a basis for the design of such experimentation. Since this experimentation is performed to make use of all the available information, in the colloquial sense, at every stage in the sequence, the usual statistical hypothesis of independence is not valid. Also, unlike classical design problems, this experimentation is devoted to estimation rather than tests of hypotheses. Many different types of practical problems will be discussed and an attempt will be made to show that the transition from one experiment to the next is a statistical problem, properly subject to mathematical study.

Information Theory and Prediction. MAX. A. WOODBURY.

A study is made of the usefulness of information theory as an aid in making and evaluating predictions. Clearly, predictions are made on the basis of information available to the predictors. Two principal problems are faced in making a prediction: First, the selection of the prediction, and second, evaluating the prediction in the light of what happens. Note that a prediction involves a selection of information available and a transformation. Naturally there is interaction between the two problems since one would change the method of prediction appropriately if the verification methods change. Here we use the measure of information proposed by Shannon and Wiener. A value is attached to probability distributions similar to entropy or negative entropy and the amount of information is measured by the change in entropy due to the additional information. In practice this involves the use of Bayes theorem. The principal predictions studied in the paper are meteorological forecasts. An ONR project has been set up to evaluate the methods proposed and with the help of the U.S. Weather Bureau

and the U.S.A.F. certain studies have been made. Here the verification evaluation consists of computing an "information ratio." This is the amount of information available in the forecast relevant to the phenomenon predicted divided by the amount necessary for a "perfect" forecast. Techniques for discarding irrelevant information have been investigated and the use of discriminant and factor analytic methods have been undertaken.

Estimation of Length of Hospital Stay from Discharge Data. CLIFFORD A. BACHRACH, *The Johns Hopkins Hospital.*

This paper considers the relationship between the two distributions of patients' length-of-hospital-stay that are commonly tabulated and used in hospital administration. These are:

1. The distribution of the lengths of stay in the hospital of patients who are discharged during a time interval,

2. The distribution of lengths of stay-to-date, of patients who are in the hospital at a given time. The event leading to a patient's inclusion in a group of patients discharged during an interval is discharge—an event which happens only once to each patient. The event leading to inclusion in a sample of the second type is presence in the hospital on a particular day—an event which may occur repeatedly to a given patient, with a frequency depending on total length-of-stay. Hence, patients who are in the hospital at a given time may be regarded as a length-biased sample of the hospital's patients; the expected mean value of the stay-to-date of such a sample is given by $(\sigma^2/2\mu)+(\mu/2)$, where μ and σ are population parameters of the distribution of lengths of stay of patients discharged during a time interval.

An analogy is drawn between a general population and a hospital patient population. The distribution of stays of patients discharged during a time interval is the analogue of the distribution of ages at death of a general population. The distribution of stays-to-date of patients in the hospital at a time is analogous to the age distribution of the general population at a census. If the stationary population model fits a hospital population, the two length-of-stay distributions are related in the same way as the d_x and l_x curves of a life table. The assumptions underlying the stationary population model are considered with reference to hospital populations of various kinds. The effects upon the model of such events as change in bed capacity and time trends in length-of-stay are pointed out. The assumptions are such that they are more likely to be satisfied in a hospital with short-term patients than in a chronic disease hospital. Data are presented to show that in one large civilian general hospital (The Johns Hopkins Hospital) the stationary population assumptions are sufficiently well satisfied to permit the use of this model for the purpose of estimating the characteristics of the distribution of the length-of-stay to-date from data on discharged patients.

Parametric Estimation of Survivorship. D. J. DAVIS, *The Rand Corporation.*

This paper deals with the application of statistical methods to the development of failure distribution hypotheses for complex mechanical systems and to the testing of these hypotheses. The two failure hypotheses considered are: *Exponential* in which non-failed systems are assumed to exhibit an equal probability of failure during equal time intervals independent of the previous operating time. *Normal* in which the probability of failure per unit time for unfailed systems increases with use in such a manner that the times-to-failures tend to cluster around the average value in a distribution similar to the normal. Statistical analysis is applied to failure data for a variety of systems involving people and/or electrical, electronic, and mechanical equipment. The raw failure data are arranged in frequency distributions which are compared to hypothetical distributions derived from either an exponential or a normal theory of failure, whichever is appropriate. The fit between the observed and hypothetical distribution is measured by means of the chi-square test. The chi-square tests indicate no significant disagreement between theory and observation in all but one of the 34 data samples examined. The 34 failure distributions are dominantly exponential rather than normal.

The Demand for Citrus Products. GEORGE M. KUZNETS, *University of California*, and RICHARD J. FOOTE, *Agricultural Marketing Service.*

The paper deals with the problems of utilizing the information collected weekly by the Market Research Corporation of America from a national panel of consumers on purchases and prices paid for various citrus items and canned juices to obtain estimates of demand elasticities. The advantages of utilizing data which exhibit variation both over households and over time are first discussed with reference to a simple model of purchase behavior. A comprehensive study is then outlined which involves separate regression analyses for summer and winter seasons extending over a recent period of 38 months in each of 10 rural-urban geographic strata. The unit of observation is a family-month and the variables include monthly purchases by panel families of each of six citrus products, prices paid for these and for four other products, family income, family size, an index of availability of frozen orange juice, and month within season. The main problem encountered in proceeding with the study is the absence of in-

formation on prices confronting the non-buying families. Probably not less than 70 per cent of the required prices for each product will have to be estimated each month. Fragmentary results from several pilot studies in progress are described These studies are designed in part to test the usability of panel data and in part to provide some indications on the basis of which a reasonable choice of a procedure of estimating missing prices could be made.

An Analysis of the Demand for Meat. AYERS BRINSER, HARRY ALLISON, and CHARLES ZWICK, *Harvard University.*

The purpose of this study of the demand for meat, fish and poultry meat items was twofold. In the first instance, it was an attempt to define certain factors affecting demand in addition to price and income and to determine the order of their relevance. The other factors include family composition, education, age, time available for meal preparation by the housewife, and ethnic group. The second objective was to test the effectiveness of a consumer panel to supply the data for such analysis. This paper is a progress report of the study outlining the design of the research and the method of data collection with a preliminary analysis of the consumer panel technique as it was applied. A full evaluation of this aspect to the study must wait until all of the data collected have been analyzed. The results of the study up to this point suggests that the consumer panel technique is a workable method for collecting data at the household level with limited research resources. It remains for the continuing analysis to show whether these data can be used to provide effective answers to such questions as whether there are differential price and income elasticities among the various subgroups in a population. On the evidence of preliminary estimates it would seem that the form and substance of the data collected will support useful quantitative analysis.

Demand Analysis from the M.S.C. Consumer Panel. G. G. QUACKENBUSH, *Michigan State College*

The M.S.C. Consumer Panel consists of about 250 families who report each week in considerable detail on their food purchases. Each family reports the quantity, price, and expenditure for each food item purchased. About 500 food items are included, and this includes all food. Each family reports its income for the week. The sample area to date has been the city of Lansing, Michigan. The project was designed to continue for 10 years. The objectives of the project are to determine price, income, and cross elasticities, as well as purchase patterns for food. The research was undertaken to help fill the need for cross-sectional analyses over time and for analysis of data of the short-period type. Examples of current research include use of single equation least squares methods to derive price elasticities for beef, pork, poultry, fish, cold cuts, individual cuts of meat (steaks, roast, hamburger, etc.), and some fats and oils. Cross-sectional analysis, using total annual purchases for individual families, has been used to derive income expenditure elasticities for food eaten away from home, food purchased for home preparation, and for some individual food products. It is concluded that the consumer purchase panel method has considerable potential as a method of securing data for demand analysis and related studies.

The Flow of Net Cash Savings through Life Insurance Companies. HARRIS LOEWY.

The flow of net cash savings through life insurance companies is made up of the savings of policy holders (internal savings) and governments, corporate debtors, and mortgagees (external savings). Internal savings equals companies' net income, and is the net of income flows such as first year's and renewal premiums, interest income, and expenditures flows such as payments to policyholders and expenses. The growing stability of internal savings is the result of the increasing relative importance of such contractual income flows as renewal premiums and interest income. Payments to policyholders will also stabilize as the result of the rapid growth of group and term insurance in force in which policyholders' equity is small. As equity declines in the future, policyholders' dissaving, a feature of the Great Depression, will also be relatively less.

Individual decisions to save are mainly incorporated in the first year's premium flow which shifts rapidly with changes in the level of income and employment. The renewal premium and interest income flows are relatively insensitive to income changes and are increasing in stability because of the contractual element they incorporate. Total income is stabilizing because of this contractual element and because the volatile first year's premium flow is becoming a declining proportion of total income. The expenditures flow is dominated by payments to policyholders. In the 30's the flow of payments to policyholders was strongly influenced by lapses and surrenders. If the decline in policyholders' equity continues as a result of the growth of term and group insurance, swings in this payments flow should be of smaller depth and amplitude. The net result of these developments should be growing stability in the annual flow of internal savings or net cash available to finance capital formation. The flows of external savings, amortization payments, are measured for 1947–1950. Added to internal savings, they show how much net cash (new money) the life insurance industry made available in those years by source to finance capital expansion.

Current Problems in Measuring Moneyflows. MORRIS A. COPELAND

Moneyflows measurements constitute a set of interlocking social accounts for the economy—sector sources and uses of funds statements. This system of social accounts is an extension of the national income and product system to show inflows and outflows for nine (instead of five) separate sectors, and to show a variety of types of transaction that do not appear in the national income and product accounts, particularly financial transactions. The way in which the economy must be divided into sectors for this purpose calls not only for replacing the business sector of the national income and product accounts with at least three financial and two nonfinancial business sectors; but several national income sectors must also be redefined. A critical current problem involved in the sectoring of the economy that is examined in some detail centers around the contrast between personal saving (disposable personal income minus personal consumption expenditure) and the net financial use of funds by households (increase in household cash and portfolios minus increase in household debts). The extent to which other regular statistical compilations (as well as the national income estimates) are currently used in making annual estimates of sector moneyflows is emphasized. Several problems involved in putting the various sources of information together into moneyflow accounts are outlined. An important but as yet unsolved problem is that of developing current quarterly estimates of moneyflows. The nature of this problem is considered in relation to the monthly and quarterly figures now available.

Estimation of the Interval Rate in Actuarial Calculations: A Critique of the Person-Years Concept. JOSEPH BERKSON, *Mayo Clinic.*

Each individual in a group for which a follow-up study is being made is observed from some defined origin of time, as for instance the time of receiving medical treatment. The observations may continue to some predesignated closing date C or be terminated by a random process. The problem is to estimate the probability of survival to time T_j. Broadly, two methods can be used, (1) If the probability of survival to T_j is given by some smooth survivorship function $P_j = F(T_j, \theta_1, \theta_2, \cdots)$ where $\theta_1, \theta_2 \cdots$ represent parameters, then with appropriate observations available, the θ's may be estimated by using some method of "fitting" the function, such as maximum likelihood, minimum X^2, least squares, etc. (2) If the function F is not known, the actuarial method is used. In this method the total period T_j is subdivided into intervals small enough so that within each interval the survivorship can be considered linear. The conditional probability of surviving to the end of the interval $[x, (x+1)]$, having survived to the beginning, is p_x, and if this is estimated for each interval, the probability of surviving to time T_j is given by $P_j = \prod_1^{j-1} p_x$. The essential problem, then, is to estimate p_x. Two estimators are discussed, (1) "Person-years" estimate, given by $\hat{q}_w = d/\text{P.Y.E}$, where d is the number of deaths observed in the interval and the "person-years exposed," $\text{P.Y.E.} = N_x - \sum_1^W (1 - T_{ix})$, where N_x is the number living at the beginning of the interval and T_{ix} is the fraction of an interval for which each of W "withdrawals" among the N_x was to be observed. (2) Maximum likelihood estimate, given by the solution of

$$q_L = d \Big/ \left[N_x - \sum_1^L \frac{(1 - T_{ix})}{1 - qT_{ix}} \right],$$

where the summation is taken over L individuals last observed living in the interval. Various approximations of each of the estimates are discussed and variances of the estimates are evaluated. It is pointed out that the "person-years" estimate implies that it is possible to define a "withdrawal" W. For the usual follow-up study a definition of W requires the imposition of a single "closing date" for all persons in the study. This involves the disadvantages that it (1) necessitates a truncation of the data that sacrifices available information, (2) makes clerical work more complicated than when the maximum likelihood estimate is used, (3) renders impossible calculation of survival rates for any series in which less than 100% of the cases are traced to the closing date.

The Problem of Within Family Contagion. WILLIAM R. GAFFEY, *University of California (Berkeley).*

The presence of contagion, or direct person-to-person transmission, in a disease process is often argued from the fact that the secondary attack rate is high compared with the primary attack rate. An example is given to show that no such inferences regarding contagion can be made from the comparison of these rates, because of the possibility of different intensities of exposure in different families. Instead, the observed variables are taken to be the times during the period of observation at which cases occur within a family, assuming families to be independent. A model is constructed which admits either (1) no contagion, (2) positive contagion (in which a case makes subsequent cases more likely), or (3) negative contagion (in which a case confers a degree of immunity on the rest of the family). The joint density of the times of occurrence referred to above is given under the hypothesis of no contagion and under the alternative of contagion (positive or negative) of a special type termed "linear."

A uniformly most powerful test of the hypothesis (1) against either (2) alone or (3) alone is found, as well as a uniformly most powerful unbiased test against both (2) and (3) together. All of these are insensitive to variations in the intensity of exposure. The first two moments of the test statistic were found as an aid to getting an approximate region of rejection when large numbers of families are involved.

Consistency of Estimators under a Specialized Bioassay Procedure. WILLIAM F. TAYLOR.

The consistency properties of various types of estimates are considered in this paper, with especial emphasis on maximum likelihood and minimum chi-square estimates. A general theorem is presented stating necessary and sufficient conditions for consistency. It is assumed that the effects of a drug are to be evaluated by giving it in varying doses to several subjects and observing the number of these subjects who respond in a given way. Let the dosage levels be $x_1, x_2, \cdots, x_s$ and the number of subjects receiving the ith dose, n_i. The proportion of subjects responding to the ith dose is called q_i. A further assumption is that the probability, p_i, of a subject responding to dose x_i is some function $f(\theta|x_i) = f_i(\theta)$, where θ is a parameter to be estimated. For several of the commonly chosen functions, f, and several methods of estimation, it has been shown that the estimates of θ so obtained are regular best asymptotically normal (RBAN) in the sense of Neyman, and thus have identical asymptotic properties.

There is more than one type of limiting situation which leads to asymptotic properties, however, and the fact that one such situation results in desirable properties may have little to do with the properties associated with other limiting situations. The results most frequently used are based on the case in which the total number of subjects is $N = \sum_{i=1}^{s} n_i$. N is allowed to increase by increasing each n_i, keeping the n_i/N and the s fixed. When this is the limiting situation considered, maximum likelihood, minimum χ^2, logit, and probit estimates are all RBAN When $N \to \infty$, not by increasing each n_i but by taking on more dosage levels, i.e., by increasing s, the above can no longer be said. This is shown with regard to the consistency of the estimates, assuming the n_i are all $\leq \overline{M} < \infty$. While maximum likelihood estimates can be shown, under certain conditions, to be consistent, under these same conditions minimum χ^2 ones are not consistent unless additional restrictions are imposed.

Use of Census Tracts in Study of Changing Residential Patterns in Metropolitan Areas. GEORGE DUGGAR.

For decision on issues of policy and theory research is needed on residential localization. Studies should be extended to encompass entire standard metropolitan areas and, in the search for a valid typology, applied with identical statistical techniques to several metropolitan areas and to several periods of residential construction. The 1950 Census of Housing permits comparison between standard metropolitan areas and between parts of areas. By comparison with the 1940 Census and by reference to data on year built new and older housing can be compared, and areas can be distinguished in terms of their rates of housing growth. Gross techniques for study of metropolitan areas are examined, centering on the 91 areas with a city of 100,000 or more and contrasting housing in their core cities and remaining area, and housing built 1940–50 and earlier. Correction of results for differences in extent of land area and use of the new urbanized area concept are found valuable. Techniques for more intensive comparative study of growth periods, of entire metropolitan areas, and of more central and more peripheral portions are examined, using census tracts and, in the untracted portions, other small statistical areas. Techniques include mapping of growth rates, growth concentration, and housing characteristics, statistical analysis of relations between these (in ways relevant to theory) and identification of small areas where growth rate or concentration suggests case study While lack of cross tabulations prevents direct comparisons between the housing in the several tracts by period built and characteristics, the general patterns of co-variation of growth rate and of housing characteristic can be examined as reflected in tract totals. It is suggested, however, that the type of distinction which has been drawn in the 1950 census between urbanized area and standard metropolitan area be drawn, also, between the urbanized and non-urbanized portions of tracts.

Theory of Behavior. EDWARD W. BARANKIN.

See *Econometrica*, Vol. 21, No. 3, July, 1953, p. 474.

Uses of Census Tracts by Housing People. DOROTHY S. MONTGOMERY.

Tract use gained impetus in Philadelphia as a result of interest in the two Real Property Surveys of the 1930's, as well as the 1930 Census. The Philadelphia Housing Association, the first housing group in the city, was among the first to use the tract for analysis and recording. As an educational, Community Chest agency, we find the tract a valuable tool in educational work as well as in our fact-finding activities. At the outset, the census tract helped to tell the story of the relationship of housing to

health, to delinquency, to dependency. Today our Association uses census data on a tract basis to interpret neighborhood housing conditions in connection with the housing tours which are an essential part of our educational program.

The technical uses of the tract are numerous, and most of them common to other communities. But I would like to describe two uses that may be unique. The first is a use that continues year after year as part of our tabulation and analysis of city permits covering new dwelling construction, conversions, and demolitions. These data have been regularly collected by our Association since 1923, and have enabled us to estimate changes in the dwelling supply throughout the city in intercensal periods with a reasonable degree of accuracy. Since 1945 these data have been tracted, which has added the important element of precise location to the previously tabulated information on type of new housing, selling price and lot size. The tracting of conversion data is of particular value since it has permitted analysis of the significant, and usually neglected changes, in the use of structures, changes that are frequently associated with the start of neighborhood blight. It may be mentioned here that the tracting of the permit data became feasible when an up-to-date street and house number directory was published for Philadelphia. Prior to the directory, tracting was almost impossible.

Tract data enabled us to discern and map the patterns in the spatial movements of Negro households in the city, and to gauge the amount of concentration and dispersion during the last census decade. Most of the influx of the decade had been absorbed in the traditional areas of Negro occupancy which had become more Negro in the process. Within these areas, however, there had been substantial shifts with the North Central Philadelphia area becoming the major center of concentration. Analysis of the tract data also showed that, contrary to general opinion, Negroes are not finding homes throughout the city; and that where expansion did take place, it was in or adjacent to areas in which Negroes already constituted a large part of the population. While over-all census figures indicated that housing occupied by Negroes was on the average of much poorer quality than housing occupied by white families, analysis of tract data revealed that there is no inevitable correlation between the race of the occupants and the quality of the housing. A comparison of two neighborhoods showed that while one, an area of high Negro concentration and severe overcrowding, has become increasingly blighted in the last two decades, the other, also an area of high Negro concentration but low density of population has shown an improvement in housing quality in the course of the racial change. This study, which could not have been undertaken without tract data, will provide the Commission on Human Relations with factual information upon which to base its educational program to end discrimination and segregation. It will also affect other aspects of the city's official policy concerning the enforcement of minimum housing laws.

Uses of Small Area Census Data in New York City. FLORENCE S. CUTTRELL. *Welfare and Health Council of New York City.*

The decade 1940–1950 brought so many shifts in the population of New York City that many organizations are studying the 1950 population and housing data for the 2563 census tracts, or the 352 larger health areas which are composed of tracts. One of the major uses by City departments, social agencies, churches and similar organizations is to estimate the needs of areas within the boroughs for housing projects, schools, health programs, case work services, recreation centers, programs to prevent delinquency, churches, etc. and to determine the location of facilities to provide these services. The City Planning Department utilizes small area data in connection with its master plan, zoning ordinances and its other responsibilities and for population forecasts. Several major research projects—studies of the aged, mental health, medical care and the teen-age narcotics problem—have used the data for selecting samples and/or for studying population characteristics and environmental factors. Utilities, commercial and savings banks, insurance companies, newspapers, construction firms and other companies use population and housing data to estimate loads or sales, to plan districts, for mortgage appraisals, to determine the need for housing or branch banks, etc. Manufacturers, advertising agencies, radio stations, and research organizations indicate dependence on block and tract data for sample marketing and public opinion surveys.

Special publications: To supplement the tract data published by the Bureau of the Census, six organizations purchased a set of the IBM summary tract cards on the 1950 population. These cards also enabled the City Department of Health and the Research Bureau of the Welfare and Health Council to publish population data for health areas. The cards also provided data for the Population and Housing volume of the "New York Market Analysis," a companion of the volume on retail trade from the 1947 Census of Business. Both volumes of this publication by the New York *Mirror*, New York *News* and New York *Times* cover 116 retail trade districts (aggregations of tracts) within the City and 21 suburban counties. The Research Bureau of the Welfare and Health Council published for census tracts and health areas: "Population of Puerto Rican Birth or Parentage, New York City: 1950."

Probability Distributions of Group Organization Theory. LEO KATZ, *Michigan State College.*

A functioning group of N individuals is represented abstractly as a multilinear, graded, directed graph on N points. Between each pair of points (individuals), the complete relationship is assumed to be analyzable into a sequence of categories, in each of which the strengths of the bonds in both directions are measurable in some scale. The hope for economy in description of the group rests on the possibility that the bulk of the information in these infinite-dimensioned vectors is contained in relatively few (perhaps, only one) of the components. A single component (one facet of the group organization), then, is represented as a directed graph and, if the relationship is all-or-none, as a simple linear directed graph. Two fundamental results provide an almost complete probability distribution theory for the classical sociometric problems. A thesis by James H. Powell of Michigan State College will contain an analysis of the structure of the sample space of sociometric investigation whidh makes possible the immediate, though not simple, expression of most of the pertinent probability distributions in terms of the numbers of graphs satisfying certain restrictions. A second result by Katz and Powell (submitted to *Proceedings of the American Mathematical Society*) gives explicitly the number of directed graphs satisfying a specific set of local restrictions

The principal classical problem not solvable as above is that of the distribution of mutual or reciprocated choices. However, an iterative procedure of application of the previous results produces the required distribution for this case. Applications of the general method include the measurement of concentration of choice, tendency toward reciprocation, development of leaders, etc. Finally, preliminary study of the bilinear graphs corresponding to two-dimensional group functions indicates that any attempt to handle these mathematically would involve invention of some new mathematics. On the other hand, this study indicates some new uses of multiple-level sociometric techniques in the field of social psychology.

Tracts in Analysis of Worker Mobility. MEREDITH B. GIVENS, *New York State Department of Labor.*

One of the uses of census tracts is to provide a basis for aggregation of data into special-purpose area groupings. Census tracts were recently utilized in this way by the research staff of the Division of Employment of the New York State Department of Labor, in a study of where people work and live in New York City and its metropolitan environs. The study was undertaken for administrative purposes, though its results are proving to be of considerable interest to all who are interested in worker mobility and the pattern of commuting in the metropolitan community. It involved comparisons, by small areas within the Metropolitan Area, of the place of work of employed persons as revealed by the Division's own employment data, compiled for employers of twelve or more workers, with the place of residence of workers living in these same areas as shown by the Federal Census of Population. Such comparisons of (1) employment at their work place with (2) Census data on employed persons counted at their homes are indicative of the "balance" or "imbalance" between the number of persons working within an area and the number of persons living in that area. The comparison does not show where the workers within a given area come from or where those who live within the area work. For this purpose a special question in the Census schedule would be required.

For purposes of analyzing and mapping data, the intra-State Metropolitan Area was divided into 57 districts. These districts represented combinations of census tracts in those localities which have been tracted, namely New York City and Westchester. For Rockland, Nassau and Suffolk, which are still untracted areas, other approaches were used for subdividing the counties. For New York City, the basic data were first prepared on a census tract basis and then combined into 210 "conversion districts"—a tentative grouping of census tracts developed with the assistance of a committee named by the local chapter of the American Statistical Association. This grouping was designed for possible use in obtaining tabulations of Census materials on a small area basis. Contiguous "conversion districts" with similarities in employment were then combined into 32 broader districts. For Westchester, the only other tracted county in the area, the 150 census tracts were combined into nine districts which conform to the developmental districts delineated by the Westchester County Department of Planning.

To obtain information on employment at place of work on a census tract basis, for combination into the broader districts, it was first necessary to code the data on insured employment according to census tract. To obtain employment data according to place of residence for combinations of census tracts, the Division purchased from the Bureau of the Census IBM summary cards showing, by tract, employment characteristics of the employed labor force for New York City and Westchester. The New York City purchase was made jointly with several other local groups using combinations of tract data for their own purposes. The task of coding the Division's employment records by tract proved difficult because of the lack of an up-to-date directory of census tracts for the City.

This project required the solicitation of one-time breakdowns of employment at each place of work from multi-establishment employers. The project demonstrates the feasibility of using census tracts as an approach to small-area analysis of data from other than Census sources in comparison with Census

data. The tract is a useful common denominator in devising larger standard areas for statistical tabulation and analysis in a metropolitan area.

Use of Census Tracts for Business Analysis. PERRY H. MYERS.

The main problem in the use of census tracts and other census data for business analysis is simply ignorance of even interested businessmen regarding census material. An informal survey suggests that about half of business executives think a census tract is a pamphlet issued by the Bureau of Census. Census tract data are essential in several major marketing developments: (1) In tracing the differentiation of the urban market and particularly the development of a new type suburban group. In this connection, it should be noted that higher incomes, the building boom, more marriages and more children have led to an accelerated movement of younger, middle income couples into the suburbs and, at the same time, concentration of lower income groups within the central parts of the larger cities. To trace this development, which has far reaching effects for both retailer and manufacturer, requires detailed study of tract data, not only in terms of population but also in terms of income, family characteristics, home ownership, etc. (2) The development of suburban retailing, particularly large department store plans for shopping centers, has required detailed analysis of local data to determine the population and characteristics of the areas to be served by these new suburban outlets. (3) Tract data are essential in consumer sampling, both in setting up a random sample and also in projecting the results to the total population of a city or to other cities. In general, it may be said that increasing purchasing power and the wider distribution of discretionary income have led to greater differentiation in the consumer market and a greater need of tracing this differentiation in terms of where people live, and who they are, through the use of the census tract.

The Relation of Census Tracts to the General Census Program. ROBERT W. BURGESS.

The activities of the Census Bureau, the methods used and the subject matters covered can be partitioned in various ways, many of them significant in directing the enterprise or proposing changes. One of the fundamental splits which effects both the determination of what shall be done and the way in which the details of various projects shall be developed is between the nation and the small area as the focus of interest. Typically an economist values statistics for the information they give of the nation as a whole and of the major economic divisions thereof. He is also interested in the changes over relatively long periods of time. The market researcher, however, and other specialists in planning economic and social activities, need information as to the population in relatively small local areas. They want to have an adequate basis for determining where a store or a school or a hospital can best be located and for judging whether one of these institutions is really meeting the present and future needs of its neighborhood. For such purposes sampling procedures do not fill the bill, and the small area statistics provided by a reasonably recent complete census are necessary.

The needs for small area statistics have thus an important bearing on determining the proper period between censuses covering a particular field. These needs also influence decisions as to the form in which census results should be stated, so as to meet the needs of users and encourage greater use of all material. It has become clear as workers in various fields apply scientific methods increasingly to the key factors of their problems that these needs for local statistics deserve careful attention. The Census Bureau, therefore, can and does cooperate with local agencies and individuals in developing census tracts and stimulating their use. The Census Bureau, however, leaves the functions of analyzing tract statistics and campaigning for educational or social projects that such analysis might suggest to individuals like subsequent speakers on this program or the organizations which they represent. Noteworthy gains are being made in the establishment of census tracts so that within a few months 46 standard metropolitan areas will be entirely divided into tracts as compared with only 11 at the time of the 1950 Census. With continued interest on the part of local committees, it is a reasonable goal that we have census tracts for all SMA's developed for the 1950 Population Census.

How the Automobile Industry Utilizes the Census Tract in Market Determination. FERDINAND F MAUSER, *Wayne University.*

The automobile manufacturers are fairly consistent in their belief that the census tract is a useful, basic statistical tool, of value in gathering distribution intelligence. Marketing strategy based upon this intelligence determines placement of dealerships. Companies consider proper placement of dealerships to be vital in their attempts to achieve maximum sales for their products. Two of the methods employed in area analysis using the census tract as the basic working framework are the standard metropolitan area approach and the central city approach. Superimposing of selected data upon area maps makes it possible to bring areas of poor dealer performance into focus. The type of data used for addition to the census tract framework varies depending upon the need and makeup on the area studied. There may be as many as 10 or 12 separate maps drawn up for a single area analysis. Statistical specifics are presented

on a three part basis: (1) the ideal situation—what the company would want if it could write the ticket exactly as it wanted, (2) the present situation as it now is, and (3) the recommended proposal in terms of specifics—the answer to the question of what should be done now. Under this system the ideal arrangement is used as the starting point. The situation as it now stands, is then related to the ideal. This brings the problems into focus. Changes then become evident and improvements are made gradually in the direction of the ideal because the ideal or goal is known.

Further Generalization of Neyman's Distributions. GEOFFREY BEALL.

It is known (Beall and Rescia, *Biometrics* 9: 354–386) that $\phi(t) = \exp. m_1 n! \sum_{s=1}^{\infty} m_2^s (e^{it}-1)^s/(n+s)!$ is the characteristic function of Neyman's contagious frequency distributions for $0 \leq n \to \infty$. Investigation now shows that n may properly assume negative values of any magnitude and thus produce an even more extensive family of frequency distributions. The distributions for $n = -1$ and $n = -2$ are intimately related to those where n is nonnegative in that, for example, the third moment steadily decreases from n infinite till $n = -2$. The distributions, however, change abruptly at $n = -3$; then the third moment becomes very great. As n decreases further the third moment again decreases and approaches the value obtaining for n very great. Presumably the Neyman distributions may now be expected to fit a great variety of data. It is known that with n small and positive they fit phenomena differing greatly from binomial situations. Current investigations suggest that much data that might be fitted by negative binomials can be as well fitted by the contagious distribution with $n \to \infty$. It may be hoped that the present extension will make them even more adequate. These further generalizations have suggested that the handling of this whole system of contagious distributions may be greatly facilitated by a transformation from the parameters m_1 and m_2 to $b = \frac{1}{2}(n+2)m_1/(n+1)$ and $c = 2m_2/(n+2)$.

The Time-Interval Approach to the Problem of Contagion. GRACE E. BATES, *Mount Holyoke College.*

The particular approach to the problem of contagion discussed in this paper is an outgrowth of a suggested treatment outlined in the last section of the paper "Contributions to the Theory of Accident Proneness, Part II", by Jerzy Neyman and Grace E. Bates, U. of Cal. Press, 1952. Starting with a stochastic model in which the probability of an individual's incurring an accident in a given time interval depends not only on the length of this interval but also, possibly, on the number of previous accidents sustained—i.e. possible contagion—the random variables considered are the time intervals from the start of the period of observation to the occurrence of each accident for an individual incurring exactly n accidents in the interval, given that the individual had already sustained exactly m accidents at the start of the period of observation. In this paper, the distribution of these time intervals was considered only in the case of contagion of a very specialized type, termed linear contagion. Specifically, this restriction on the type of contagion provides, for example, that each additional increase in the number of accidents previously sustained (from 3 to 4, or from 10 to 11, for example) yields the same percentage increase or decrease in that individual's chances of avoiding accidents in the period of observation. Under this restrictive condition the distribution of time intervals is independent of the *number* of previous accidents sustained and the time-interval data needed for testing the hypothesis of no contagion does not require this information.

For this model, uniformly most powerful tests of the hypothesis of no contagion against either of the one-sided alternatives and a uniformly most powerful unbiassed test for the two-sided alternative case are obtained. The statistic used for these tests is the grand mean of all the time intervals for all the individuals. The exact power function can be given explicit form but is so cumbersome to apply that a method of approximating the power is outlined.

Contemporary Topics in Statistical Physics. G. W. PRESTON, *Philco Corporation.*

The history of the development of the classical theory of statistical mechanics is briefly traced. It is recalled that the behavior of macroscopic quantities of matter can be correctly described by the use of the Laws of Motion and the theory of probability. Not only the condition for thermodynamic equilibrium, but also the equations for the rate of processes, can be given by probabilistic statements. The fundamental statistical attributes of matter are the relative independence of the fundamental units of matter and their very frequent, though proportionately brief and violent interactions. The fact that the fundamental units of matter are nearly independent greatly simplifies the distribution problems, whereas their frequent interactions guarantee that the system will assume a very large number of microscopic configurations during the course of any thermodynamic measurement. The fundamental results of the recent theory of fluctuations are shown to suggest a possible approach to a statistical mechanical theory of non-equilibrium thermodynamics. Finally, the necessity of including in the expression for the physical entropy of a system the entropy of information obtained about the system by experimentation is shown to imply an equivalence between the second law of thermodynamics and certain basic theorems in the theory of statistical estimation.

Statistical Problems in Physics. MARTIN J. BERGER.

This paper reviews certain aspects of the relationship between physics and statistics. Some reasons are brought out why physicists in the past have paid comparatively little attention to problems of statistical inference. The point is made that this neglect is unjustifiable and deplorable. Various branches of statistics are then examined from the point of view of their importance to physics. Finally, some typical problems of statistical inference are described that arise in the interpretation of physical experiments. It is shown that there is a large class of problems in physics to which the existing statistical methods can be readily applied with great advantage, while others of a more specialized nature would require further development of statistical theory along novel lines. The hope is expressed that physicists will begin to make greater use of existing theory, and that statisticians will look into new problems raised in physics to the mutual advantage of both sciences

The Statistician in a Research and Development Laboratory. B. B. DAY, *U. S. Naval Engineering Experiment Station, Annapolis, Maryland.*

Using a particular research and development laboratory as a case history, the steps followed to put a statistician to work therein are outlined in detail. Three major points are developed: (a) the selling job required at the different levels with illustrative material presented; (b) the organizational set-up in the Laboratory for most effective work; and (c) the statistician on the job—qualifications, working relations, and responsibilities. Some details of the internal operation of a Statistical Office are indicated. The paper concludes with a discussion of what the future holds for the statistician in the Laboratory.

Survey of the Theory of Finite Sampling. JOSEPH F. DALY, *U. S. Bureau of the Census.*

Not many years ago the design of sample surveys depended heavily for its success on the reputations of its practitioners, who strove mightily to "validate" their samples with the aid of collateral information. Today, notwithstanding some popular impressions to the contrary, finite population sampling is one of the most respectable and thoroughly practical applications of mathematical statistics. The present quite satisfactory state of the theory of sampling finite populations is based on two main characteristics of the problems which the theory is designed to solve. In the first place, it is possible in many of these problems to devise methods of sample selection and estimation such that the resulting sample estimates can be expected to obey the laws of probability. The amount of variability in the estimates arising from the fact that they are based on a sample rather than on a complete count can therefore be measured by the same techniques and with the same precision that one can measure the variability in successive drawings from a table of random numbers. In the second place, it is frequently possible to devise formulas which will serve as good approximations to the way in which the costs of projected sample surveys will vary with the size of sample and with the manner of selection of sampling units. This makes it possible to define objectively the notion of an "efficient" or an "optimum" survey design, namely one which minimizes the variability subject to fixed total cost.

Important work has been done in the way of devising techniques of sample selection which are more efficient than simple random sampling (e.g. selection with unequal probabilities, controlled random selection across strata, etc.) and on developing estimation formulas which make maximum use of available collateral information (e.g. regression estimates on one or more variables). One unsatisfactory aspect of the present theory is that it refuses to evaluate sample designs which are not subject to the laws of probability. Recent work on statistical decision theory based on the theory of games promises to shed some light on the properties of probability sampling methods in relation to larger classes of sample designs, including the selection of samples based on expert judgment. It is now known, for example, that under certain conditions simple random sampling represents a minimax strategy. Further developments along this line can be expected in the next few years.

Some Finite Sampling Concepts in Experimental Statistics. JEROME CORNFIELD, *National Institutes of Health.*

It is necessary to know the expectation of each of the mean squares in an analysis of variance in order (a) to choose a proper error term (b) to estimate components of variance. These expectations depend on what was sampled. Usually this dependence is expressed by denoting each observation as a linear compound of fixed and random variables, making certain assumptions about these components, and deriving the expectations. Different assumptions lead to different expectations, however, and in any analysis more complex than the one-way classification it is not always possible to choose unequivocally among the different possible assumptions.

A general way out of this difficulty is provided by finite sampling concepts. This is illustrated in the two-way classification, for which one can assume a population of elements classified into R rows and C columns with N elements in each of the RC cells. A sample of r rows, c columns and n elements within the rc cells in the sample is taken. If we assume the sampling is such that each row has the

same probability of selection (r/R), that each column has probability of selection c/C, and each element n/N, that the probabilities of sampling any two rows, columns, or elements within cells are respectively $r(r-1)/R(R-1)$, $c(c-1)/C(C-1)$ and $n(n-1)/N(N-1)$, and that rows, columns and elements within cells are sampled independently, the expectations can be derived without further (and more disputable) assumptions by elementary (but laborious) algebra to be

Sample mean square for	Expectation
Rows	$\sigma_e^2 \frac{N-n}{N} + n\frac{C-c}{C}\sigma_I^2 + nc\sigma_R^2$
Columns	$\sigma_e^2 \frac{N-n}{N} + n\frac{R-r}{R}\sigma_I^2 + nr\sigma_c^2$
Interaction	$\sigma_e^2 \frac{N-n}{N} + n\sigma_I^2$
Error	σ_e^2

where σ_e^2 is the within cell mean square for the population of RCN elements, $N\sigma_I^2$ the interaction mean square for this population, $NC\sigma_R^2$ the row mean square and $NR\sigma_c^2$ the column mean square. The expectations for Eisenhart's model I, model II, and mixed model follow as special cases.

STATISTICAL ABSTRACTS

All communications concerning this section should be addressed to the Abstracts Editor, Professor George E. Nicholson, Jr., Chairman of the Department of Statistics, University of North Carolina, Chapel Hill, North Carolina.

Anis, A. A., and Lloyd, E. H., "On the range of partial sums of a finite number of independent normal variates," *Biometrika*, 40 (1953), 35–42.

Given a random sample of n, $\{X_1\}$ from a normal population. The partial sums, S_1, S_2, $\cdots$, are formed, where $r=1, 2, \cdots, n$, $S_r=\sum_{i=1}^{r} X_i$. The average value of the range of the n partial sums is shown to be $\sqrt{2/\pi}\sum_1^{n-1} r^{-1/2}$. R. L. ANDERSON, *North Carolina State College.*

Anscombe, F. J., "Sequential estimation," *Journal of the Royal Statistical Society*, 15 (1953), 1.

The author reviews the literature on sequential estimation, presents some approximate methods for solving a number of particular problems and discusses the possible usefulness of sequential estimation procedures.

A sequential estimation procedure is a procedure in which the number of observations is not fixed in advance, but depends, according to some definite rule, upon the observations themselves. An example of such an estimation procedure, due to Haldane, is the estimation of a binomial proportion p when it is desired that the standard error of such an estimate be roughly proportional to p. The rule is to take observations until the number of successes X is equal to some prearranged number $X \geqq 2$. Then an unbiased estimate of p is $p=(X-1)/(N-1)$ where N is the total number of observations. An unbiased estimate of σ_p^2 is $p(1-p)/(N-2)$. Another example is that of estimating the mean of a normal population with unknown variance by a confidence interval of width d and confidence coefficient $1-\alpha$. Stein has given a double sampling procedure in which a first sample of fixed size n_0 is taken and then a further sample of $N-n_0$ observations is taken where N depends on the observations in the first sample. Reference is made to work of Girshick *et al.* for obtaining an estimate of the unknown proportion p at the end of a sequential test on a binomial population. The author asserts that estimation formulas valid for fixed sample sizes are asymptotically valid for sequential sampling if the sample size is large. He develops a sequential estimation procedure for estimating the mean of a normal population with unknown variance by a confidence interval of width d and confidence coefficient $1-\alpha$. If t_α is the normal deviate which has probability $\alpha/2$ of being exceeded, the rule is to stop taking observations when s^2 the unbiased estimate of the variance is first less than or equal to $l^2n^2/4t_\alpha^2$. Modifications of this rule are discussed which reduce the error of the estimate. It is shown that the expected sample size corresponding to this procedure is approximately $4\sigma^2t_\alpha^2/l^2+1+t_{\alpha/2}^2$. Similar results are obtained for estimating the difference between two means, and estimating the birth and death rates in a simple birth-death process. Finally it is conjectured that perhaps the main practical benefit to be derived from studying the properties of sequential stopping rules is to assess the degree to which the usual fixed sample size estimation formulas are affected. GEORGE E. NICHOLSON, JR., *University of North Carolina.*

Altman, Irving B., "Relationship between sample size and AOQL for attribute single sampling plans," *Industrial Quality Control* (January 1954), 29–30.

A chart showing the AOQL values for a large variety of single sampling plans is presented as a function of sample size and acceptance number. GERALD J. LIEBERMAN, *Stanford University.*

Bailey, Norman T. J., "The total size of a general stochastic epidemic," *Biometrika*, 40 (1953), 177–85.

The author presents a stochastic treatment of the problem of estimating the number (w) who will become infected if one infected person is introduced into a group of n susceptible individuals. Diagrams are presented of the probability of w being infected for $n=10$, 20 and 40 and $\rho=n/4$, $n/2$ and n, where ρ is the ratio between the rate of removing infected people from the group and the infection rate. The average w is also computed. The results are extended to the distribution of w for house-

holds, when n is small; formulas are given for $n = 1(1)5$. The author summarizes his results as follows: "The model used here is not adequate for diseases with short infection periods, such as measles, but its adequacy for other infections requires testing." R. L. ANDERSON, *North Carolina State College.*

Bartlett, M. S., "Approximate confidence intervals," *Biometrika*, 40 (1953), 12–19.

The likelihood derivative, $T = \partial L/\partial\theta$, is used to derive approximate confidence intervals for various θ, assuming T is normal. An approximate correction for skewness (κ_1) of T is introduced, using $T_\lambda \equiv T + \lambda(T^2 - I)$, where I is $\sigma^2(T)$. λ is approximated as $-\kappa_3/6I^2$. Also if $\kappa_3 = 0$, a method is given to adjust for kurtosis. Three examples are presented. R. L. ANDERSON, *North Carolina State College.*

Cohen, A. C., "Estimating parameters in truncated Pearson frequency distributions without resort to higher moments," *Biometrika*, 40 (1953), 50–7.

Using the method of moments, estimating equations are obtained which involve only the first four sample moments in contrast to the first six previously employed. The general case of doubly truncated four parameter distributions and special cases of doubly and singly truncated samples are worked out or their solutions indicated. It is pointed out that these equations can either be solved directly or by an iterative process. An example is worked out to illustrate the method and afford an evaluation of the technique. D. C. HURST, *North Carolina State College.*

Cox, D. R., and W. L. Smith, "The superposition of several strictly periodic sequences of events," *Biometrika*, 40 (1953), 1–11.

N sources, each producing events at regular intervals but with unequal periods, θ_i, $i = 1, 2, \cdots, N$, are pooled. This produces a sequence of unequal intervals. The frequency distribution for these intervals is derived. There is a point frequency, Q, for the largest of these intervals, which is the same as the smallest of the θ_i. A measure of the interval variability, displayed in a variance-time curve, can be used to distinguish the given sequence from a random sequence. Approximate procedures are given for large N.

Finally given a pooled sequence, a method is available to estimate the number of sources, N.

Three examples are given. This reviewer could not check the results for Q. R. L. ANDERSON, *North Carolina State College.*

Craig, C. C., "Some remarks concerning the Lot Plot Plan," *Industrial Quality Control* (September 1953), 41.

The author presents a non-technical evaluation of Dorian Shanin's Lot Plot Plan. He discusses three main points:

1. The misuse of lot plot for process control.
2. The severity of the operating characteristic curve of the Lot Plot.
3. The dependence on histograms. GERALD J. LIEBERMAN, *Stanford University.*

Craig, C. C., "Note on the use of fixed number of defectives and variable sampling sizes in sampling by attributes," *Industrial Quality Control* (May 1953), 83–86.

Under the assumption that samples are taken with replacements from small lots or that lots are large enough so samples taken without replacement will behave effectively as if they had been taken with replacements (infinite lots), the author examines the problem of attribute sampling where the random variable is the number of items sampled, n, before a fixed number, m, of defectives are found. A control chart based on the random variables n/m or n is presented. GERALD J. LIEBERMAN, *Stanford University.*

Craig, C. C., "On the utilization of marked specimens in estimating populations of flying insects," *Biometrika*, 40 (1953), 170–76.

An observer catches butterflies, marks them and then releases them. This is repeated s times in a given area and the numbers of butterflies caught one, two, ... times are recorded. Five methods are used to estimate the total number (n) of butterflies in the area, three assuming a truncated Poisson distribution and two a special distribution due to Stevens. Approximate variances are obtained for each estimate.

Two worked out examples are included, plus the results of 14 other catches. The author uses χ^2 to test the agreement between actual and expected frequencies and finds 15 significance probabilities $>50\%$ and the 16th $>30\%$. R. L. ANDERSON, *North Carolina State College.*

Evans, D. A., "Experimental evidence concerning contagious distributions in ecology," *Biometrika*, 40 (1953), 186–211.

The frequencies for a variety of counts made on English and American plant species, insects and larvae and moth eggs are

fitted to three different theoretical contagious distributions: negative binomial, Polya-Aeppli and Neyman Type A. The mathematical formulae for fitting are included. For each distribution, two methods of computing are compared: (i) use of sample mean and variance, (ii) use of sample mean and proportion of zeros. A method of selecting between (i) and (ii) is presented. The Neyman Type A seemed to be best for the English plants, the negative binomial for the insects and moth eggs, and none of the three seemed to be satisfactory for the American plants.

Agreement between actual and expected distributions is tested by two statistics originally advanced by Anscombe. The usual χ^2-test for frequency data is also used, but this reviewer notes a tendency to pool too many classes on the tails. The Anscombe statistics seem to be more sensitive to aberrations from theory than is the all-purpose χ^2-test. R. L. ANDERSON, *North Carolina State College.*

Fisher, Walter D., "On a pooling problem from the statistical decision viewpoint," *Econometrica*, 21 (1953), 567–85.

The problem posed is that of finding a logical basis for grouping a set of K random variables into a set of G subgroups ($G \leq K$). A procedure is outlined which places this problem in the framework of statistical decision theory. With a proper specification of the random variable and with the explicit recognition of additional assumptions and conditions, the logical subgrouping is determined from the observed sample data.

The following formulation of the general problem is offered. A sample $(x_1, \cdots, x_k)$ consists of one observation drawn independently at random from each of K disjoint "cells." The random variable x_i has expectation θ_i, variance σ_i^2, and a measure of importance π_i. The θ_i are unknown, the σ_i^2 and π_i are known positive numbers. The problem is to arrange the set of K cells into G mutually exclusive subgroups according to some partition P, and to choose a decision vector $t_P = (t_{1P}, \cdots, t_{KP})$, where t_{iP} denotes a decision number or "estimated characteristic" of that group to which the cell i is assigned by partition P (t_{iP} for cells i assigned to the same group will be identical). Conditions may be imposed limiting the number of subgroups G and also the partitioning procedure. There remains a number of theoretically admissible alternative partitions and a number of possible decision vectors for each partition. A choice of the optimum decision vector t^* involves the multiple choice of an optimum partition P^* and an optimum set of decision numbers $t_{iP}{}^*$. Following decision theory a loss function $W(\theta, t_P)$ is defined which measures the cost incurred by making the decision t_P when the true parameter point is θ. The vector t^* is determined as the set of decision numbers which minimizes the expected value of W over the domain of possible samples.

The procedure is formulated in specific terms with reference to the following hypothetical problem: A public price control agency is assumed to have the task of setting selling prices for a given commodity. The commodity is sold in K independent markets, each with a linear demand function of the form $p_i = 2\theta_i - (1/\pi_i)q_i$, where p_i is price, q_i is quantity sold, and π_i is a positive number, $(i = 1, \cdots, K)$. It is assumed that the price control agency in setting prices is guided by two conflicting policies: (1) to be "fair" to sellers, and (2) to attain low administrative costs. Administrative costs are assumed known and are assumed to increase with the number of different prices set by the agency. A measure of "unfairness" to sellers is assumed in the present problem to be represented by the departure of actual realized returns under adminis ered prices from maximum realizable returns under free market prices. Under the conditions assumed the expression for the loss function becomes $W(\theta, t) = c_G + \sum_{i=1}^{K} \pi_i (t_{iP} - \theta_i)^2$, where the first term on the right represents total administrative costs incurred if G different prices are set and the second term represents aggregate returns sacrificed by sellers as a result of the administered prices t_{iP}.

Apart from the price agency illustration but assuming a loss function of the same specific form, stochastic properties are ascribed to the set θ and a risk function (expected value of the loss) defined. A limiting Bayes solution to the decision problem is derived (and presented in an appendix) under the assumption of a sequence of a priori distributions of θ having constant density and increasing range. It develops that the optimum decision procedure t^* is the following: For a sample x (in the price agency problem x_i is estimated equilibrium price in market i) select for each value of G (number of different administered prices) the admissible partition $P_G{}^*$ of the K cells (independent markets) which minimizes the quantity $S_P = \sum_{i=1}^{K} \pi_i (x_i - \bar{x}_{iP})^2$, where $\bar{x}_{iP}$ is a weighted sample group mean. The weighted means of the sample observations within groups formed by the optimum partition for a specified G give the optimum set $t_G{}^*$. Denote the above sum of squares

corresponding to P^* for a specified G by S_G^*. Then the optimum set t^* is given by G which minimizes $(c_G + S_G^*)$.

This solution is applied to sample data on retail prices of fresh tomatoes in nineteen independent retail markets in California. In this numerical example restrictions are placed on the partitioning procedure which limit the number of admissible partitions for each G greater than 1, and G itself, limiting the maximum number of different administered prices to 4. In addition, a schedule of administrative costs c_G is assumed for $G = 1, \cdots, 4$, and π_i is assumed uniform for all markets. The resulting optimum set of administered prices is presented. It turns out that the optimum number of different prices in this case is 4, the maximum allowable.

In a concluding part the author offers some comments with respect to limitations and possible extensions of the proposed pooling procedure and discusses briefly the relation of the pooling problem to certain other subjects such as the problem of aggregation discriminant analysis, classification problems, and factor analysis. IVAN M. LEE, *University of California.*

Fox, Karl A., "A spatial equilibrium model of the livestock-feed economy in the United States," *Econometrica*, 21 (1953), 547–66.

In this paper the author obtains numerical solutions for a ten-region spatial equilibrium model for feed grain of the type formulated by P. A. Samuelson in "Spatial Price Equilibrium and Linear Programming," *The American Economic Review*, June, 1952. The problem is not formulated explicitly in a *statistical estimation* framework, although actual data and approximate relations derived from actual data are employed in the computations. The paper serves primarily to illustrate unique numerical solutions to the spatial equilibrium model for a particular simplified formation of the problem and under alternative assumptions regarding the magnitudes of certain variables entering the model.

Quantitative results showing regional equilibrium prices, net trade, and regional patterns of trade for conditions with respect to freight rates, feed supplies, livestock numbers, and livestock prices approximating those for 1949–50 are shown. Solutions are also shown for two alternative sets of conditions: (1) Conditions like 1949–50 except that freight rates are assumed uniformly 50 per cent higher and (2) conditions like 1949–50 except that feed supplies are assumed lower in the Corn Belt and Northern Plains regions by 30 per cent and 75 per cent, respectively. An "application to forecasting" is also illustrated, which differs from the problem solved for 1949–50 conditions only in that July, 1947 forecasts and December, 1947 estimates of feed grain production are used in arriving at regional feed grain supplies. IVAN M. LEE, *University of California.*

Gulde, Harold J., "Acceptance sampling by variables using the range," *Industrial Quality Control* (November 1953), 18–25.

Formulas and charts are presented for constructing variables sampling plans using the range and average range. The results are based upon the Normal approximations. GERALD J. LIEBERMAN, *Stanford University.*

Horsnell, G., "The effect of unequal group variances on the F-test for the homogeneity of group means," *Biometrika*, 40 (1953), 128–36.

The author extends the 1951 investigations of David and Johnson (*Biometrika*, 38: 43), using four groups in a single classification analysis of variance. Most of the computations are based on three equal within-group variances with a fourth variance three times as large. Assuming equal group means, the actual significance probabilities corresponding to nominal 5% and 1% significance levels of F are obtained for equal and various combinations of unequal numbers of observations per group.

The power of the usual F-test to detect one divergent group mean (out of the four) is approximated both when the variable group has and has not the divergent mean.

The author summarizes his results as follows: If there is no very clear information as to heterogeneity, use equal group frequencies. If it is known that one group is more variable, make sure that this group does not have less observations than the others; if possible, take a few extras in this group.

An alternative to the usual F-test is also mentioned. R. L. ANDERSON, *North Caroline State College.*

Kamat, A. R., "On the mean successive difference and its ratio to the root mean square," *Biometrika*, 40 (1953), 116–27.

Given a sequence of n normal variates $\{X_i\}$ with means $\{\mu_i\}$ and common variance σ^2. The mean successive difference is

$$d = \sum_{i=1}^{n-1} \frac{|x_i - x_{i+1}|}{n-1}.$$

The first four moments of d/σ are derived when all $\mu_i = \mu$. The standard deviation, β_1 and β_2 values of d/σ are tabulated for $n = 3$ (1) 10(5) 30, 40, 50; a type I Pearson curve was used to obtain approximate upper and lower $\frac{1}{2}$, 1, $2\frac{1}{2}$ and 5% percentage points. Exact results are available for $n = 3$. Empirical sampling was used to check some results.

The usefulness of d over the standard deviation, s, is shown when the μ_i shift.

The distribution of $W = d/s$ is considered when all $\mu_i = \mu$; its mean, standard deviation, β_1 and β_2 values are given for $n = 5(5)30$, 40, 50. Approximate percentage points are also presented. R. L. ANDERSON, *North Carolina State College.*

Kamat, A. R., "Incomplete and absolute moments of the multivariate normal distribution with some applications," *Biometrika*, 40 (1953), 20–34.

Given n normal variates, $\{x_i\}$, with zero means and given variances and covariances. Absolute moments are moments of $\{|x_i|\}$. Exact results are given for $n = 2$, 3, and power-series for derivations when $n > 3$. These results are applied to the distribution of $T' \equiv \Sigma |x_i|$, using Pearson-curve approximations. R. L. ANDERSON, *North Carolina State College.*

King, E. P., "Estimating the standard deviation of a normal population," *Industrial Quality Control* (September 1953), 30–33.

The efficiency of various estimates of the standard deviation of a normal population is presented when nk observations are grouped into k subgroups of n observations. GERALD J. LIEBERMAN, *Stanford University.*

Latscha, R., "Tests of significance in a 2×2 contingency table: Extension of Finney's table," *Biometrika*, 40 (1953), 74–86.

Finney's 1948 table (*Biometrika*, 35: 148) for making exact tests of independence in a 2×2 contingency table with fixed border totals is extended through $A = 20$. This table is of the form

	Number of		Total
	Successes	Failures	
Series I	a	$A-a$	A
Series II	b	$B-b$	B
Total	$r=a+b$	$A+B-r$	$A+B$

The experimenter defines a success and Series I such that $A \geqq B$ and $a/A \geqq b/B$. For a given A, all combinations of B and a are included which will give a significant result, where the significance probabilities are $\alpha = .05$, .025, .01 and .005 (one-tailed). For each combination of A, B and a, the author gives the value of b ($< a$) which is just significant at the α-level, b_α, and the exact probability that $b \leqq b_\alpha$, for a given A, B and r. R. L. ANDERSON, *North Carolina State College.*

Lindley, D. V., "Estimation of a functional relationship," *Biometrika*, 40 (1953), 47–49.

It is desired to estimate α and β in the linear model $V = \alpha + \beta U$, where the only observations on U and V are subject to error and all distributions are normal. U and V are allowed to be either mathematical variables or random variables. If the observations on one of the variables are at predetermined fixed points, the so-called "controlled" observations, it is demonstrated that the estimating equations for α and β can be obtained in the usual way. This result also applies if these "controlled" observations are allowed to be random. If there is no "control" on either set of observations, the estimation is deemed impossible without recourse to supplementary information. D. C. HURST, *North Carolina State College.*

Moore, P. G., "A sequential test for randomness," *Biometrika*, 40 (1953), 111–15.

Given a sequence of observations, where each observation falls into one of two alternative categories. A sequential procedure is presented for testing the hypothesis, H_0, that these observations occur in random order, against the alternative H_1, that there is dependence of the kind found in a simple Markoff chain. The procedure is illustrated with annual rainfall data. C. E. GATES, *North Carolina State College.*

Scheffé, Henry, "A method for judging all contrasts in the analysis of variance," *Biometrika*, 40 (1953), 87–104.

The author presents and proves the validity of a method for making further inferences about the contrasts among a set of true means or main effects $\mu_1, \mu_2, \cdots, \mu_k$, following the rejection of the hypothesis $H: \mu_1 = \mu_2 = \cdots = \mu_k$ by the conventional F-test with $k-1$ and ν degrees of freedom in the analysis of variance. The method is based on a probability statement concerning the infinite totality of contrasts of the type

$$\theta = \sum_{i=1}^{k} c_i \mu_i$$

where the c_i are any known constants, satisfying the condition

$$\sum_{i=1}^{k} c_i = 0.$$

If the assumptions usual in the analysis of variance hold and $\hat{\theta}$ and $\widehat{\sigma_{\hat{\theta}}^2}$ denote the estimates of θ and the variance of $\hat{\theta}$, then it is proved the probability is $1-\alpha$ that the values θ of *all* possible contrasts simultaneously satisfy

$$\hat{\theta} - S\widehat{\sigma_{\hat{\theta}}} \leqq \theta \leqq \hat{\theta} + S\widehat{\sigma_{\hat{\theta}}} \qquad (1)$$

where S^2 is $k-1$ times the upper α point of the F-distribution with $k-1$ and ν degrees of freedom. The result holds for any values of the unknown parameters. The result may be used for the interval estimation of any contrast of interest or to declare any estimated contrast "significantly different from zero" or not, according as the corresponding interval (1) excludes $\theta=0$ or not, contrasts suggested by the observed means $\hat{\mu}_i$ included.

A method has been prescribed by Tukey for the special case where all the $\hat{\mu}_i$ have the same variance and all pairs $\hat{\mu}_i$, $\hat{\mu}_j$ $(i \neq j)$ have the same covariance, and where there is interest in only a subset of the totality of contrasts, namely the $(1/2)k(k-1)$ differences $\mu_i - \mu_j$. This latter method is shown to have greater sensitivity in this situation than the method based on (1) in the sense that the confidence intervals are shorter.

The author examines the operating characteristic of the method by determining (i) the probability that all contrasts whose true values are zero would be declared "not significantly different from zero," and (ii) the probability that all normalized contrasts whose true values are greater in absolute value than some specified bound will be declared "significantly different from zero." D. G. HORVITZ, *North Carolina State College.*

Stevens, W. L., "Tables of the angular transformation," *Biometrika*, 40 (1953), 70–73.

The angular transformation is defined as

$$\theta = 50 - \sqrt{1000}\ \text{arc sin}\ (1 - 2p),$$

where p is the estimated proportion for a binomial population. Values of θ are given to 3 decimal places for

$$p = 0(.0001).02(.001).5,$$

and for selected proper fractions. R. L. ANDERSON, *North Carolina State College.*

Stuart, A., "The estimation and comparison of strengths of association in contingency tables," *Biometrika*, 40 (1953), 105–10.

A modification of Kendall's rank correlation coefficient, t_c, is proposed for measuring strength of association between two characteristics in an $r \times s$ contingency table. Conservatively approximate confidence limits are set for the population association. An approximate test is provided for the difference between the coefficients calculated by this method for two $r \times s$ contingency tables. The coefficient used is

$$t_c = \frac{n-1}{n} \cdot \frac{m}{m-1} t_a$$

where t_a designates Kendall's coefficient of rank correlation. S. WEINER, *North Carolina State College.*

Tanner, J. C., "A problem of interference between two queues," *Biometrika*, 40 (1953), 58–69.

Imagine a bridge of certain length AB only wide enough to admit a single lane of traffic, with traffic arriving at both ends. Vehicles V_1 arrive at A at random and at an average rate q_1 per unit time. A particular V_1 takes time α_1 to cross AB, but may enter AB only provided no V_2 vehicle is in AB, and if no other V_1 vehicle has entered AB during the previous time β_1. Similar sets of conditions, α_2 and β_2 affect the V_2 vehicles arriving at B. The mean delay time $\bar{w}_1$ and $\bar{w}_2$ for $\alpha_1 > \beta_1$ and $\alpha_2 > \beta_2$ is obtained as a function of t_1 and t_2, the time for a block of V_1 or V_2 vehicles to cross, and r_1 and r_2, the number of vehicles waiting to enter the bridge. Explicit solutions are made available for the special cases $\beta_2 = 0$ and $\beta_1 = \beta_2 = 0$. Further applications to the problems of delays in crossing or entering a heavily travelled artery are described, and delays due to intersecting lines of traffic discussed. Two short tables are included, giving the values of $\bar{w}_1$ for $\beta_1 = \beta_2 = 0$, $\alpha_1 = \alpha_2 = 1$ and for $\bar{w}_2$ when $\alpha_2 = \beta_2 = 0$, $\alpha_1 = 1$. J. S. HUNTER, *North Carolina State College.*

Whitfield, J. W., "The distribution of total rank value for one particular object in m rankings of n objects," *The British Journal of Statistical Psychology* 6, part 1 (1953) 35–40.

The title clearly indicates a class of problems not appropriately handled by Kendall's treatment of the general m ranking problem. As an illustration one can consider a situation in experimental social psychology where one person in a group is instructed to play a predetermined role and each of

the other members of the group is asked to rank his fellow members (including the experimental person) in connection with certain characteristics. Under the assumption that for the first ranking each value (1, 2, $\cdots$, n) has equal probability the frequency distribution for m rankings of n objects is constructed; the mean total rank value is $\frac{1}{2}\ m(n-1)$, its variance is $[m(n^2-1)]/(12)$, $\beta_2=3-(6)/(5m)-(12)/[5m(n^2-1)]$. Tables of exact probability values are given for m and n up to 8. For higher values of m or n $[|\text{Total Rank Value}-\frac{1}{2}\ m(n-1)|-(\frac{1}{2})]/\sqrt{[m(n^2-1)]/(12)}$ is approximately a normal deviate. For example, if eight judges rank seven objects, and the experimental object has a total rank value of 43; the approximation gives $P=.0318$ against the exact value of .03113. HERBERT SOLOMON, *Columbia University*

BOOK REVIEWS

Introduction to the Theory of Statistics. *Victor Goedicke.* New York: Harper and Brothers, 1953. Pp. xii, 286. $4.50.

BERNARD L. WELCH, *University of Leeds, England*

THE present text sets out the fundamentals of statistical method in a sound and readable manner and can be recommended to students who are starting upon the subject from the very beginning.

Since no use is made of the calculus the author is often forced to be purely descriptive and to dispense with much theoretical justification. This seems to be inescapable when only a small prior knowledge of mathematics is assumed. The reader is not even required beforehand to be familiar with the use of logarithms and with simple graphs. Sections are devoted to these topics at appropriate points as they are needed. I feel, myself, that writers of elementary books on statistics should not be expected to go so far as this in remedying deficiencies in the mathematical equipment of their readers. If a student has not already at least some facility with graphs and logarithms then he is, I believe, ill-advised to start to grapple with the theory of statistics at all.

The decision of the author to include sections on permutations and elementary probability, however, is not open to the same criticism and his introduction to these topics is admirable. In his treatment of frequency distributions, the expression of frequency as a proportion per t unit (where $t=(x-\bar{x})/\sigma$) leads on naturally to the standard normal curve as a graduation of frequency distributions arising in many different practical contexts. Some theoretical justification of the general use of the normal curve is attempted by including a numerical investigation of the limiting behavior of the binomial distribution. This section, however, makes more difficult reading than the rest of the book and could possibly be simplified in places.

The treatment of correlation is excellent, the emphasis rightly being placed on the square of the correlation coefficient rather than r, itself—the idea being to stress that the important factor is the fraction of the variance accounted for by a straight-line relation. Multiple correlation is dealt with again from the same angle and the examples used to illustrate it are well-chosen.

The introduction to the sampling variability of such simple statistics as a sum, a difference, and a mean value is clear and leads on to the general problem of testing statistical hypotheses where, however, some statements are made which are open to criticism. For instance in an inquiry where the effects of two drugs on the period of convalescence after illness are being compared, the author gives an example where the probability is 0.0018 that a difference between drug-means greater than the one observed could have arisen on the null hypothesis. He goes on to deduce that "the available evidence indicates that the probability that the drug does *not* affect the duration of convalescence is only 0.0018, while the probability that it does affect duration is 0.9982." Such a statement as this is certainly unwarranted and some-

thing more by way of explanation seems to be called for at this point. One must grant, of course, that it is somewhat difficult to explain in an elementary fashion what the interpretation of the figure 0.0018 should be. The devotion of more space to the clarification of this general problem of testing hypotheses would not be a waste, however, for as they stand, the sections on this topic suffer by reason of compression.

The concluding section of the book draws attention to certain commonsense viewpoints which the novice is apt to overlook.

The standard of printing is high throughout and altogether this is a useful addition to the texts already available on the elementary parts of statistical theory.

Statistical Inference. *Helen M. Walker* and *Joseph Lev.* New York: Henry Holt and Company, 1953. Pp. xi, 510. $6.25.

PALMER O. JOHNSON, *University of Minnesota*

THIS text is written primarily for non-mathematical students. The complete text is planned for a course two-semesters in length. The problems connected with writing text-books for the type of students specified are particularly difficult if the responsibilities of the author(s) are to be fulfilled. There is the problem of what mathematics can be assumed on the part of the reader. A second problem resides in the selection of content. Since texts of the kind discussed here are written for those who are to become practitioners or research workers, the most important statistical procedures must be described and the computational problems arising must be discussed. The methods must also be illustrated with representative examples from the field or fields of application, usually with the particular field with which the author is familiar. There is likewise the selection of problems to be solved by the student for testing his understanding of the theory and methods. The writer of texts for non-mathematical students is under particular obligation to describe rigorously and to indicate in his own illustrations the assumptions to be fulfilled if the procedures are to be validly applied. This is particularly important in a book whose readers lack the knowledge necessary to check the author's statements.

Unfortunately there are no properly designed experiments to test the efficacy of text-books on modern statistical methods. It is, therefore, difficult to make criticisms that are completely objective. Most statisticians have strong convictions not only on the relative importance of various statistical procedures but also on how mathematical (even for non-mathematical students) a text-book should be.

In the reviewer's judgment, the authors of this text have in general succeeded in writing a text that reaches a high level of attainment in meeting the criteria suggested above.

The mathematical requirements would be chiefly met by a good command

of arithmetical and algebraic processes. There is considerable use of probability symbolism, multiple subscripts, multiple summation, and specified limits of summation. In this connection, well stated explanations and practice exercises in the processes involved are provided. With respect to previous statistical training, the student with good ability and familiarity with symbolism can acquire the elementary calculation processes, such as calculation of the mean, standard deviation, regression, and correlation from the treatment in the text.

Skilful use is made of intuitive reasoning and of effective graphical devices for developing a functional understanding of statistical concepts and processes underlying statistical inference.

The nature of statistical inquiry is introduced in Chapter 1 by a simple experiment. Chapter 2 lays the basis for a more formal treatment of probability and its role in statistical inference. Fundamental concepts are introduced, such as variable, population, probability, random observation, probability distribution, independence, elementary laws of probability, parameter, and statistic. Particular application is made to the binomial distribution. Chapter 3 extends the concepts and methods of statistical inference to the types of problems encountered in the use of the binomial model. An unusually clear exposition is given of such fundamentals as sampling distributions, critical regions, levels of significance, point and interval estimation, tests of alternative hypotheses, power of a test, one- and two-tailed tests, and the two types of errors. Turning to Chapter 4 an extension is made of the previously developed concepts from the binomial to the multinomial populations comprised of several discrete classes. The content deals mainly with chi-square, its sampling distribution, properties, the chi-square curves and tables. A number of applications are made, emphasizing calculations and interpretation. No mention is made of the criticisms and limitations of the chi-square test of "goodness of fit," particularly that the power of the test to detect disagreement between hypotheses and observations is determined largely by the size of the samples.

In Chapter 5 and in subsequent chapters consideration is given to problems in inference involving continuous variables. Certain general concepts underlying distribution of a continuous variable are developed. Notions of the parent population, unbiased estimates, and a test of normality are included. Chapter 6, on sampling distributions, presents a number of empirical distributions of commonly used statistics. The frequency distributions of various statistics are plotted and the mathematical curves superimposed. Consideration is given to statistics following the chi-square, student's t, and the F-distributions. The property of independence of statistics is well portrayed by plotting the joint frequency distribution of the means and standard deviations of samples from a normal population. Chapter 7 presents methods for testing hypotheses and for estimating population means and differences between means of two populations. The test of alternative hypotheses is

especially well treated including the graphical interpretation of the power of a test. Limited consideration is given to the design of samples in surveys. Estimates for the mean and variances are given for stratified and cluster samples. Along similar lines Chapter 8 considers the problems of inference concerning variances and standard deviations of normal populations. Special consideration is given to the use of the tables of the F-distribution. Bartlett's and Hartley's tests for the equality of variances of several samples are applied to practical problems. There is also a test for differences between variances of correlated measures.

Chapter 9 presents some of the simpler applications of the analysis of variance. A k-sample problem of testing the equality of means is illustrated in an experiment designed to determine if order of shooting three different ranges in archery had any effect. The underlying mathematical model is specified including assumptions. In this connection it would have been desirable to carry out the tests of normality and of equality of variances rather than merely to indicate that these assumptions had been satisfied. The test of the equality of means is carried out by estimating the unknown population variance from the variation among the three sample means and from the variation among the archers within groups, then by taking the ratio of the two estimates, the variance ratio, or F, as the test statistic. This approach makes it difficult for students to perceive why these sums of squares were selected. The approach through stating the mathematical model in algebraic form or in the form of a linear hypothesis seems to be more informative. It would entail the use of the method of least squares or of maximum likelihood to obtain the appropriate sums of squares. In this chapter there is an excellent explanation of the form of the F-distribution, including curves of the sampling distribution of F and $1/F$.

In Chapter 10 methods of statistical inference are applied to bivariate data. A mathematical model for linear regression and the sampling distributions of statistics in linear regression are presented as the bases for testing hypotheses and for interval estimation of parameters. The bivariate population model and the normal bivariate population model are treated. The distribution of the product moment correlation coefficient is described. Tests of hypothesis about ρ are carried out and confidence intervals for ρ set up. There is another chapter (Chapter 11) given to other measures of relationship. The chief contribution here is a test of significance for point-biserial-r and a brief summary of Tate's unpublished study comparing bi-serial and point-biserial-r.

Chapter 12 is devoted to the statistics of measurement. This chapter should serve to bring to the attention of the research worker the importance of errors of measurement particularly in tests of significance and in estimation. Most of the content of this chapter is of long standing and indicates how little workers in this field have been influenced by the developments in modern statistical inference. Most of the old algebraic formulations are

reported in this chapter. Perhaps an opportunity would have been afforded to report some beginnings that have been made and other help that could be had from ideas and developments in statistical inference. Examples are Wilk's L_{mvc} criterion for parallel tests, Votaw's test of compound symmetry, component mean square analyses in regression problems, the practical sampling problems of reliability and validity, and the uses that are beginning to be made of information theory in mental test theory and construction.

Chapter 13 gives a lucid explanation of the methods and interpretation of multiple regression and correlation. Illustrations are given of the Doolittle method and Fisher's modification of the Doolittle method in obtaining a multiple regression equation and tests of significance. There is no consideration in this or other chapters of other methods of multivariate analyses. Thus the discriminant function designed for obtaining the best system of weights of various independent variables for distinguishing among groups of individuals is not treated, nor are such methods discussed as Hotelling's T for testing the significance of the difference between means of multivariate normal populations, and methods of classification of an individual based on a number of measurements into one of several categories. Examples of the latter methods are the generalized distance function and the statistical decision function.

In Chapter 14, the analysis of variance is extended to the analysis of two or more variables of classification. This chapter bases its discussion around data derived from studies employing modern experimental designs. In following through the analysis of the designs presented, the student should gain a realistic conception of the role of statistics in the analysis of the data from modern designs. It would have been desirable to have given more explanation to the role of statistics in the planning stage of the design and greater emphasis to the functions of randomization and replication.

A few points of criticism will be made to an otherwise admirable chapter:

> On p. 349, μ in H_r and H_e appears to be the same value, which in general is not true unless the value is zero. On p. 354, the test of the hypothesis that the mean of the 4 diagonal cells is equal to the mean of the 12 non-diagonal cells by means of the t-test is not valid unless probability corrections are made. The latest proposals for the solution of this problem are the analyses suggested by Scheffé and by Tukey. There is no explanation in the text of the origin of expected mean squares in tables 14.6 and 14.12. Nor is any statement made tying these values up with Model II of the analysis of variance. On p. 363, recognition is given to the circumstance that when several measures are obtained from the same individual, the measures are likely to be correlated. It is not clearly indicated how, if at all, the effects of the correlated observations have been removed.

Chapter 15 gives special consideration to the use of the analysis of covariance in group comparisons on one variable when information is available on another or on several other variables which are correlated with it. Skilful demonstration is given by the analysis of data from a number of important research problems in this area.

In Chapter 16, certain special uses of percentiles are illustrated in analysis

leading to inferences. In Chapter 17 several useful transformations are discussed. The final chapter discusses and illustrates the use of a number of non-parametric methods in testing hypothesis and in setting up confidence intervals.

The book has an exceptionally large number of useful tables and charts (20 in all), besides a table of squares, square roots, and reciprocals, a table of four-place logarithms, and one of random numbers.

There is a five-page glossary of symbols. A detailed table of contents and a well-prepared index facilitate the location of items of interest. All but four of the chapters contain a list of references.

Eleven of the chapters contain exercises for the student. They are most frequently interspersed within the chapter. These exercises are well designed to test understandings of terms, phrases, principles, symbols, formulas, use of tables, etc. Only three chapters contain exercises dealing with real data. To the reviewer, it seems that in courses for practitioners abundant opportunity should be given to deal with real data—fictitious data do not often give experiences that will be met with in research. There are no such opportunities in Chapters 14, 15, 17, 18, for example, which deal with the types of situations most often encountered in practice or likely to be used by students in their own research. These types of situations are likely to be most motivating and to develop an appreciation of the dynamic qualities of statistics.

In summary, the reviewer believes that this text will make a valuable contribution by elevating the plane of instruction in applied statistics. While the fields of application are chiefly in education and psychology, the book could profitably be used as a basic text in other fields. It will particularly appeal to instructors in the social science fields because so many of the current texts are written for students in biology and in agriculture.

Statistical Methods in Experimentation: An Introduction. *Oliver L. Lacey.* New York: The Macmillan Company, 1953. Pp. xi, 249. $4.50.

OSCAR KEMPTHORNE, *Iowa State College*

THIS text is designed for use in a one-semester course in general statistics as applied to experimentation, and assumes one semester of college algebra as a background. The subject matter of the book is indicated by the chapter headings: The aim and problems of statistics in experimentation; Experimental design; Interpretation; Probability (I)—the probability of discrete events; Probability (II)—probability in a continuum; Three chapters on the normal distribution; Tests of significance of means and differences between means; Enumeration data; Correlation; Regression; Fiducial limits; Experimental design; Appendix of tables.

The plan of each chapter is to give first a general discussion of the matter under consideration, then examples, which are in the form of questions, and

answers to these examples, and finally problems for the student. The general form of the book, the level of discussion of topics, and the question-and-answer technique strike this reviewer as being very good and very appropriate to a first course in statistics for students with no mathematical background. The format and general style are very clear and should not be a source of confusion or annoyance to the intended readers.

A number of criticisms can, however, be made.

With respect to the design of simple comparative experiments, the discussion is somewhat naive. The author states that there are four general methods of controlling variation due to extraneous factors (1) elimination (2) equalization, (3) balancing, (4) randomization. The first two are never completely realizable, and the third is only realizable with special assumptions which may be very arbitrary. The basic fact is that the first three methods are valuable adjuncts to randomization and it is rare that randomization can be dispensed with. In any case if the physical act of randomization is not done the validity of the statistical inference is questionable as Fisher originally stated. The combination of system and randomness in a design is little discussed, and this is perhaps the most important feature experimentally.

The discussion of the size of experiment is made generally in terms of "minimal adequacy," although the word efficiency is used occasionally. It would seem desirable pedagogically to conform to the usual use of the word "efficiency," and to retain the notion of power instead of introducing a new notion.

It would be desirable that all statistical texts make the distinction between regression relationships and functional relationships. Originally regression was a relationship in a multivariate population, but now is widely used for the relationship of y to x, where x is not a random variable. When one uses regression analysis on the growth curve of plants, one is really estimating a functional relation of growth (y) to age (x) and in this case there is no point in considering an underlying population of x values from which one has a sample. In a true multivariate situation one may be interested in correlation or regression, but in the functional relationship situation the correlation has no real significance. All the problems in the chapter on regression are the functional type. Uses of regression in multivariate populations could be discussed more fully, and examples given.

The author acknowledges the stimulation of R. A. Fisher's *The Design of Experiments*, but certain passages indicate that there is some confusion in the author's mind on some basic questions. For example he gives as a sample question, "Does home economics training result in a better chance for marriage than arts college training?" And then gives the answer "This may be attacked experimentally by keeping comparative records for a number of years of home economics and arts college graduates. Obviously care would be needed to make groups truly comparable in respects other than academic training." The reviewer disagrees entirely with this answer; the answer that

is given is not an answer to the question posed. The question posed can be answered strictly by giving girls home economics training or arts college training at random, without regard to their inclinations or anything else and then observing their marital success. The answer which would be obtained by the stated procedure is the answer to the question "Do home economics graduates have a better chance for marriage than liberal arts graduates?" The lack of appreciation of the difference between the question posed and the question answered appears to be rather widespread in the social sciences. Also there is a question about the answer: what is meant by "truly comparable?" The adjective "truly" appears to have taken on some mystic significance with some scientists and the use is better avoided. In addition the reviewer does not like the use of the word "experimentally" in the answer, in that one is merely observing an uncontrolled situation. Nor does the reviewer favor the use of "identical" in statements such as "we should present them (*two drinks*), therefore, in identical glasses" p. 3 (cf. the use of "equal" on page 18).

Writing really good elementary texts in statistics is a very difficult job. The present book is a good attempt, and would be more than good if the defects mentioned above were corrected.

Methods of Statistical Analysis in Economics and Business. *Edward E. Lewis.* Boston: Houghton Mifflin Company, 1953. Pp. viii, 686. $5.50.

Z. Szatrowski, *University of Buffalo*

In the opinion of the reviewer, Professor Lewis has written a definitely good book on statistical methods in business and economics. He has done this by adding to the conventional material some of the fundamental concepts of modern statistical analysis. His discriminating selection of subject material and careful allocation of space according to the importance of basic topics enables him to cover more and get it across. The exposition (text, illustrations, and problems) is excellent.

The organization of the book is reasonably conventional and departures from the pattern of the past are improvements. There are 686 pages, five parts and an appendix. Part I (pages 1–96), the introduction, discusses the purpose and problems of statistical methods. Also this section includes an explanation of graphs, tables, and "Statistical Numbers and the Problem of Accuracy." Part II (pages 97–186) presents descriptive statistics for frequency distributions, i.e., measures of central tendency including the geometric mean, measures of dispersion, skewness, and kurtosis. The application of these descriptive statistics is explained clearly. The illustrations involving comparisons are very effective. In connection with this section, the reviewer would like to call attention to an inaccuracy in the specification of class limits and mid-points. The author's presentation results in a bias. To illustrate, on page 67, Table 6, for weekly earnings recorded to two decimals,

the first class is given as 22.00–23.99 with mid-point 23.00. More accurately it should be 21.995–23.995 with mid-point 22.995.

Part III (pages 187–298) is entitled "Statistical Inference." Here are contained various tests of significance, for small as well as large samples, including the application of the F and χ^2 Test. This section concludes with a discussion of quality control methods. In the reviewer's opinion, this is an outstanding presentation. The problem of inference is given proper emphasis. The various applications, which in so many texts are scattered throughout the book, now appear in one section. The author succeeds in covering a great deal in this part because he has a logical unified organization of the material and because he develops each well-chosen illustration of inference just far enough to present the idea clearly.

Part IV (pages 299–370) very adequately discusses the subject of index numbers, their construction, the problems, and current applications. Part V (pages 371–486) deals with time series analysis, trends, cycles, and seasonals, including the changing seasonal. This section is not involved but thorough. In connection with trend determination, the author includes a brief discussion of the use of transformations, so widely applicable in non-linear trend problems. Part VI (pages 487–649) is a clear presentation of correlation, multiple as well as simple, including tests of significance.

The Appendix in this book contains a list of references, formulas, and tables of square roots and logarithms. Tables of the Normal, t, F, and χ^2 distributions appear in the text where applications are discussed.

The book contains an adequate number of problems which appear throughout the chapters to illustrate specific topics which the author has just discussed. In general, the problems are well chosen and involve the minimum of calculation. The format of the book is in keeping with the high quality of the contents.

An instructor using this book can expect his students to get much out of the text. Also he can expect many of his students to become interested in statistics. This book should enjoy success as a text. In addition, since it is easy to understand and relatively complete and up-to-date in its coverage, it should serve as a very useful reference book in business and economic statistics.

Sampling Inspection by Variables. *A. H. Bowker* and *H. P. Goode.* New York: McGraw-Hill, 1952. Pp. xi, 216. $5.00.

H. C. HAMAKER, *Philips Research Laboratories, Eindhoven, Holland*

THE first thing one notices about a new book usually is its title, and it is therefore important that this title should convey a concise and correct impression of the content of the book it covers. The title of the book under review does not quite satisfy this criterion and I am afraid that many statisticians who ordered the book because of its attractive title, will feel slightly

deceived. For it does not bring a general and comprehensive survey of "*sampling by variables*" and the variety of problems involved; the book mainly describes "*a particular system of sampling by variables*" to be used where an attribute sampling plan according to the tables of *Sampling Inspection* by the Statistical Research Group, Columbia University, or the Military Standard 105A is to be replaced by a variables sampling plan.

The basic pattern of the sampling plans described is that given by Wallis in Chapter 1 of the Statistical Research Group's *Techniques of Statistical Analysis*. The quality characteristic of the products inspected is supposed to be measurable on a variables basis; an upper or lower limit, U or L, is set, and we wish to control the percentage of items with a quality beyond one of these limits. To this end we measure a sample of n items, compute the average $\bar{x}$ and the standard deviation s from these measurements, and then require

$$\bar{x} + ks < U \text{ or } \bar{x} - ks > L, \tag{1}$$

as the case may be, k being a constant to be chosen in relation to practical requirements.

This single sampling technique is now supplemented with double sampling plans which are specified by three constants k_a, k_r, and k_t, and two sample sizes, n_1 and n_2. When after the first sample

$$\bar{x}_1 + k_a s_1 < U, \text{ the lot is accepted;} \tag{2}$$

and when

$$\bar{x}_1 + k_r s_1 > U, \text{ the lot is rejected.}$$

Since $k_r < k_a$ these conditions cannot be satisfied simultaneously. If neither of conditions (2) is satisfied, that is when

$$\bar{x}_1 + k_r s_1 < U < \bar{x}_1 + k_a s_1, \tag{3}$$

we proceed to take a second sample, the requirement being that ultimately

$$\bar{x}_t + k_t s_t < U \text{ for acceptance} \tag{4}$$

when $\bar{x}_t$ and s_t are average and standard deviation computed from the pooled measurements of first and second sample. Sequential sampling according to a similar principle would become complex and inconvenient in practice, and has not been included for that reason.

The above criteria hold if we assume the standard deviation σ of the lot unknown and variable from lot to lot. If past experience has shown that σ does not vary from lot to lot, known-sigma plans may be used by replacing s by σ in inequalities (1) to (4) and modifying the constants k.

Likewise a separate chapter is devoted to cases where we have to consider simultaneously an upper and a lower limit.

The choice of a sampling plan according to these principles is facilitated by an extensive set of tables and charts, which have been arranged after the

model set by *Sampling Inspection* of the Statistical Research Group and its offspring the Military Standard 105A.

After a choice has been made between three inspection levels, the plan is determined by lot-size and AQL-class. A full set of charts with operating characteristics of all sampling plans has been included, while additional tables are provided for finding the AOQL, or LTPD, and for deciding on tightened or reduced inspection.

The text accompanying these tables gives full instructions as to their meaning and application. One chapter of 17 pages dealing with the theory requires considerable mathematical knowledge for its understanding, but all other chapters have been written in a simple and clear style, while mathematical formulas have been largely avoided.

Yet despite the absence of formulas the text is mainly theoretical in concept, and many of the puzzles encountered in practice are not even mentioned as the following examples may illustrate.

The limits U or L, see above, are taken for granted, whereas it may often be an important preliminary step to decide by a statistical investigation whether they have been fixed in a reasonable manner.

It is repeatedly stressed that a first step is to decide what is the item to be inspected. In most cases, however, this is obvious enough, but it is more difficult to decide how the measurements are to be performed. For instance, in the case of metal sheets (pages 86–93) the thickness and Rockwell hardness may easily vary from the centre to the border of the sheets and if so it may be difficult to decide how and where the measurements have to be performed. Similarly special measures may sometimes be required to avoid systematic differences between different inspectors or apparatus. Such points have not been considered.

The sample size for known-sigma plans is about one-fourth that for unknown-sigma plans. One of the most useful functions of a control chart used in combination with sampling by variables would therefore be to indicate when sigma is sufficiently under control for the use of known-sigma plans; but chapter 9 provides no guide on this point.

In practice, lot sizes may vary considerably and sample sizes with them. But a control chart for varying sample sizes loses a great deal of its attractive simplicity and, therefore, of its technical importance. Hence if we are to derive full profit of the control chart technique in combination with sampling by variables a constant sample size may be of tremendous advantage. We should not forget the Lot-Plot technique, which has taught us the practical value of a constant sample size, and a simplified technique for computing the standard deviation.

On two points of methodology I disagree with the authors.

The standard deviation of a sample is defined by

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{(n - 1)}},$$

using $n-1$ in the denominator instead of n as is more common. This, I think, is a happy change; in more advanced applications of statistics the concept of the number of degrees of freedom is so fundamental that it should be introduced at an early stage, if we are not to be led into considerable confusion and contradiction later on. But the standard error of s as defined above is

$$\sigma_s = \sigma/\sqrt{2(n-1)}$$

and not $\sigma/\sqrt{2n}$ as stated on page 107; the factors B_3 and B_4 for control limits in table J should be corrected accordingly.

More serious objections can, I believe, be raised against the techniques recommended on pages 62 and 63 for finding the process average. Let us suppose that the critical limit U is 20, and that we receive 19 lots with average $\mu=17$ and standard deviation $\sigma=1$, and one lot with $\mu=25$, $\sigma=1$; 19 lots are completely good, 1 is completely bad, the process average is 5%. But if we pool these 20 lots into one grand lot, this grand lot will have an average $\mu=17.4$ and a standard deviation $\sigma=2.0$.

Should we consider this grand lot to be normally distributed, we would conclude that the process average per cent defective is 9.7%, a gross overestimate. It might be objected that nobody would ever consent to treat such a heterogeneous mixture of lots as one lot with a normal distribution, but this is exactly what is recommended on the pages cited. Samples drawn from successive lots are pooled into one grand sample, and the mean and standard deviation of this grand sample are treated as if they represent a grand lot with a normal distribution. This is only correct if the lots are under control, but then sampling inspection is not needed. Sampling inspection fundamentally applies where there is not sufficient control; in that case within-lot and between-lot variation should be clearly distinguished, and pooling of data from different lots is not permissible. The correct procedure is to estimate the percentage defective for each lot separately and average these percentages.

Summing up them, where Military Standard 105A or the Statistical Research Group's *Sampling Inspection* for attributes is to be replaced by sampling by variables, the reader will find in this book an immediate guide; while he who wishes to develop a system of his own, will find in it a great amount of information, which will be useful if combined with practical experience and ingenuity.

An Introduction to Statistical Science in Agriculture. *D. J. Finney.* New York: John Wiley and Sons, 1953. Pp. 179. $3.75.

H. W. NORTON, *University of Illinois*

THIS little book has chapters entitled "The need for statistics," "Some problems of rates and frequencies," "Probability," "Properties and uses of distributions," "An experiment to compare two varieties," "The reduction of error," "Factorial design," "Sampling," and "Correlation and regression."

In addition, there is a preface, a valediction, references, and index. Its object is "to impart a knowledge of the general principles on which statistical science is founded and of the manner in which it enters into so many agricultural problems . . . the emphasis throughout is on the principles illustrated by the examples rather than on the arithmetic and algebra of calculation." I am one of that group of statisticians (mentioned by Finney) who think the correlation coefficient is now of very trifling importance, especially in an introductory book. This topic can hardly have much of the merit of familiarity attributed to it by the author, and the space devoted to it might better have been used to give a simple example of analysis of covariance.

A number of formulas are given and used arithmetically, but there is no algebra nor other mathematics with the sole exception of the derivation of a formula for the sample number required to realize a specific confidence interval. Type, paper, and binding are good and there are few typographical errors.

It is a curious fact that there appears to be a real need for nonmathematical books on statistics. That is, it is curious that there are many people who are established scientists, and many others offering to become scientists, who have so little mathematics (in spite of a great need thereof) as to warrant the preparation of severely nonmathematical books expounding an essentially mathematical subject. It is reassuring in a way, in the midst of current attacks on education in the public schools, accompanied sometimes by assertions that standards are higher elsewhere, to note that this book is based on (and presumably is the substance of) a "course of eight lectures . . . to undergraduates reading agriculture at Oxford." I have long held the view that the research scientist's most urgent statistical need is for an appreciation of the essential ideas of the design and analysis of experiments, and that those ideas can be appreciated without recourse to mathematics. Therefore I expect Finney's book will probably have wide appeal, and it seems well suited to its purpose of conveying appreciation of principles.

An important feature of the book is its stress on the relation between the research worker and the statistician, and the continual emphasis on the desirability of expert statistical advice at all stages of the work but especially before experimentation is begun. This culminates in the "Valediction," with its twelve rules of respectable statistical conduct for research workers. The first is "When you propose to undertake an experiment or a sampling investigation of a kind that presents any novelty to you, consult a statistician at an early stage of your planning." This will serve as an example while enabling me to point out my only substantial criticism of these rules, that the phrase "of a kind that presents any novelty to you" should have been omitted on the ground that recognition of relevant novelty will often require consideration by an experienced statistician. I was glad to find (p. 126) Finney's observation, "even a year . . . is not an excessive margin, for consideration

of the relative merits of different designs takes a long time and is better done in short spells over a period than in a single concentrated effort." How often the statistician has occasion to agree!

There are a few statements to which exception must be taken, and which do not appear to be merely the result of such simplification (usually called *over*simplification) as reasonably would be expected in such a book. The assertion (p. 25) that Yates' "correction is always $\frac{1}{2}$; and (p. 33) "applies *only* to 2×2 tables" may be defensible in a chapter which emphasizes contingency tables, but is likely to prove confusing to the novice who looks into more than this one book. The statement (p. 156) that the covariance "must always be intermediate in magnitude between" the two variances is wrong: it may not be larger in magnitude than the square root of the product of the variances. The statement (p. 102) that "both the smaller dressing of sulphate of ammonia and the ammonium humate gave yields significantly higher than that without nitrogen" may prove disconcerting to the careful reader, because the differences mentioned are exactly at the 5% point (following Finney's practices in rounding) for the comparison of yields in pounds per plot, but do not quite reach the 5% point for tons per acre, and retention of additional decimals leads to the conclusion that significance is reached only by accumulation of rounding errors. Occasionally the choice of words is inappropriate as (p. 75) the use of "indubitable" to describe the superiority of one variety to another after their difference of yield has proved significant. Some statements would be improved by minor changes, such as (p. 60) "conditions that are as alike as possible," which is likely at once to coincide with the readers' prejudice and to miss an opportunity to emphasize the potential importance of a proper distribution of effort in experimentation, would better say "as practicable" or "as reasonably possible," and (p. 68) "the difference is greater than would occur by chance" needs to say "would be at all likely," so as to keep always before the novice the idea that *any* difference may have occurred purely by chance. I could not verify the probability near the top of p. 57, nor the fiducial limits on p. 159. There is a misprint in table 6.1, where the mean for block V should be 19.0; on page 98, where the divisor in the formula for F should be 3.0; and on p. 168 the references are to table 9.3.

Agricultural Prices. *Frederick Lundy Thomsen* and *Richard Jay Foote.* New York: McGraw-Hill Book Company, 1952. Pp. xi, 509. $6.50.

L. J. NORTON, *University of Illinois*

THIS is a thorough revision of a 1936 book by the senior author. It is a text book intended for students principally in agricultural colleges. It bears the imprint of extended experience in teaching, government economic

research, and practical experience in business. It is divided into three parts:

1. A review of simple economic principles, discussion of some materials, presentation of the elements of economic fluctuations, and an analysis of government price programs under the over-all heading of "price determination and discovery."

2. Price analysis and forecasting.

3. A series of commodity reviews, labeled "commodity prices." The emphasis is on simple, direct tools rather than on complex mathematical methods.

Economic statisticians, particularly those engaged in commodity analysis, will find Section II of particular value. The chapters on commodity prices are not up to the standard of those in the first two sections in getting the problems into clear focus.

The authors hold a hard-headed point of view with reference to the value of overly elaborate analysis in forecasting. Their point of view might be summarized: Get all the information you can out of the best analysis you know how to make. Test your results with all the logic and common sense you have. Apply them with judgment and be on the alert for facts not used in your basic analysis. From his observations on forecasters and forecasts this reviewer deems this to be a highly sensible point of view. The authors suggest that pin-point forecasts are not needed but rather "a degree of interpretation which will enable the user to anticipate the direction and to some degree the extent of the movement" (p. 351).

In discussing price determination both the aggregate approach and the older individual commodity approach are used. No time is spent on hair-splitting differences or too-elaborate types of analysis. The students should get the points. Perhaps the terms supply and demand are used too broadly and so lack precision, but this may be justified on the grounds of teachability. The discussion of the "relation between cash and future prices" is not up to the standard of the book. Quite possibly many traders in the future's markets have better bases for formulating expectations than this treatment would lead one to believe. In some form or other many of them have grasped some of the ideas which the authors later develop as to reasons for price trends. In discussing the general price level too much emphasis is put on banking and credit. Perhaps the authors do not infer causation but their treatment would leave such an impression in the reader's mind. After a period when monetary inflations, devaluations, and many other developments have affected the levels of prices in various parts of the world, their analysis seems oversimplified. This may be important because in connection with the general price level the price analyst may find an important element which he may have omitted from his statistical calculations. In this area the authors do not seem to have done work in depth as in other phases of price analysis.

On the whole this is a good job and both students and workers in the field of price analysis will find it valuable.

A Textbook of Econometrics. *Lawrence R. Klein.* Evanston, Illinois, and White Plains, New York: Row, Peterson and Company, 1953. Pp. ix, 355.

KENNETH J. ARROW, *Stanford University*

THE appearance of this book marks an important stage in the development of econometrics, since for the first time there is available a useful textbook. To say that Klein's work is the best of its kind would be correct but very inadequate, since he has virtually no competition. Some other books, which have appeared under similar titles, are very useful in their own right, but they do not meet the need of a beginning course in econometrics (as opposed to mathematical statistics) presupposing a reasonable but not excessive knowledge of mathematics as well, of course, as a good knowledge of economics.

Econometricians are very fortunate indeed that this text, which will undoubtedly be the standard one for many years, is so truly excellent. Klein's preeminence as a practitioner of the econometric art is well known, but one would not necessarily expect such an extraordinary level of didactic skill from a mathematically-inclined economist (although the readers of *The Keynesian Revolution* might not be surprised).

After an introductory chapter on the econometric approach, there is a forty-page summary of the basic principles of mathematical statistics. While about as good a job has been done as is possible in the space,[1] I seriously doubt the usefulness of the attempt. The material the author seeks to cover requires, I would judge, about one semester to master, and no attempt to speed up the learning process is apt to be successful. I think it preferable to require a basic course in mathematical statistics as a prerequisite.

Chapter III, entitled, "Estimation of Aggregative Models," is in many respects the core of the book. A simplified macro-economic model is developed, and the method of least squares is then introduced as the maximum-likelihood method of estimating one of the equations. In more general cases, it is noted that the problem of identification arises; the concept is rigorously defined and its application to linear models stated. The estimation of an aggregative model by the method of maximum likelihood applied to the whole system is then considered, with attention paid to the special cases where the statistical method has a simple form. The method of instrumental variables and the limited information method are next introduced, though the explicit derivation of the formulas from maximum likelihood considerations is omitted. The chapter concludes with a discussion of confidence intervals for the estimated parameters.

The next chapter is a detailed description of the computational designs of the least squares, maximum likelihood full-information (under the assumption of a diagonal covariance matrix of the disturbances), and maximum likelihood limited-information methods. The exposition is first-rate; for a student, it could hardly be improved upon.

[1] Some minor errors in this chapter have been noted in the review by R. Solow, *American Economic Review*, Volume 43 (1953), 947–50.

Chapter V deals with some special problems arising in the analysis of particular sectors of the economy. The discussion is largely carried on by examples, and a number of ingenious devices for linking aggregative and disaggregated models within the framework of maximum likelihood methods are discussed. These are followed by a brief but excellent discussion of input-output analysis from the econometric viewpoint, and by a somewhat fragmentary but highly suggestive study of cross-section analysis and its interpretation, particularly in relation to the problem of aggregation. Chapter VI is rather miscellaneous; it deals principally with the problem of prediction from an econometric model. On one statistical point I think Klein is wrong; an interval forecast is not a tolerance interval in the ordinary sense but something more like a confidence interval. That is, we say simply that the probability is .95 (or whatever figure we wish) that the next observation will fall within our interval, rather than that the probability is .95 that at least .80 of all future observations (with the same values of the predetermined variables) fall within our interval. Otherwise, however, the discussion is excellent and brings together a good deal of otherwise inaccessible material. The last chapter deals with some special problems, such as variables measured with error, weighted regressions from sample survey data, and seasonal variation and related topics. Again, the quality of the discussion is admirable and the material well worth presenting.

Since this text will unquestionably be the standard one in the field, some critical advice to the instructor is called for. I believe it will be useful to supplement the present book with some readings in the recent Cowles Commission monograph edited by W. C. Hood and T. C. Koopmans,[2] and even in the earlier Cowles Commission monograph edited by T. C. Koopmans.[3] Some of the particular points I have in mind follow:

1. Although the general definition of identification is excellent, the special case of linear models is treated rather quickly (though correctly), although it is the only one used in practice. The student's appreciation would, I feel, be enhanced by the examples so well worked out by Koopmans in Monograph 14, Chapter II. The proof of the identification criterion is rather simple and would not be out of place in an econometrics course at the level of the book under review; such a proof is given by Koopmans and Hood in Monograph 14, Chapter VI.

2. Perhaps the greatest weakness is the cursory handling of the distinction between exogenous and endogenous variables. The special role of exogenous variables is absolutely essential in simultaneous equations estimation; but Klein gives no explicit definition in terms of the probability distributions and only a vague verbal definition. His discussion at some points seems to

[2] W. C. Hood and T. C. Koopmans, eds., *Studies in Econometric Method*, Cowles Commission Monograph 14, New York: John Wiley and Sons, 1953, referred to below as Monograph 14

[3] T. C. Koopmans, ed., *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10, New York: John Wiley and Sons, 1950, referred to below as Monograph 10. This book has a not entirely undeserved reputation for incomprehensibility, but on some points no better statements have yet appeared in print.

imply exogenous variables are simply those which belong to the province of other sciences than economics,[4] as though the proper method of analyzing phenomena were dependent on the way we divided up their study. Formally correct definitions are given in Monograph 14 in Chapter I by J. Marschak and in Chapter VI by Koopmans and Hood, but there is little concrete discussion of the meaning of exogeneity even there. The only really satisfactory treatment is that of Koopmans in Chapter XVII of Monograph 10, and this applies only to time series. The concept of an exogenous variable in cross-section analysis is tricky, though the formal definition is clear enough, and it may also be noted that, unlike the case of time series analysis, lagged endogenous variables are not necessarily predetermined.

3. The flexibility of the limited-information method in handling estimation problems which do not conform to the strict conditions for which it was originally designed is not given enough stress. The handling of nonlinear systems by making the system linear in a new set of variables which are nonlinear combinations of the original set is not anywhere stated as a general principle, though an example discussed on pp. 191–196 makes use of the method. Estimation when some of the predetermined variables not in the equation being estimated are unobservable and the use of the limited-information method in a good many cases where some variables are subject to errors of observation are two other topics worthy of mention in a course of this nature. The material is very well summarized by Chernoff and Rubin in Chapter VII of Monograph 10.

4. The study of econometrics should really include such topics as index numbers and budget studies. For example, the use of Engel curves to derive functional index numbers, a subject to which Klein himself has made such an important contribution, would hardly be taken up anywhere else in the economics curriculum.

These remarks should not be taken as more than advice to the user of the book. Some of the additional material mentioned above was not available to the author at the time of writing of the book. The minor blemishes noted should not obscure the fundamentally excellent organization, full coverage, realistic flavor, and clarity of presentation of the book under review.

Bibliography on Income and Wealth, Volume II, 1948–1949. *Phyllis Deane*, editor. Cambridge: Bowes and Bowes, 1953. 37s. 6d.

DOROTHY S. BRADY, *Bureau of Labor Statistics*

THE annotated list of publications on national income and wealth, social accounts, national budgeting, and related subjects in the *Bibliography on Income and Wealth, Volume II* was compiled on essentially the same criteria of selection and classification as Volume I. The references included in the Bibliographies have been limited on the whole to the literature relating to

[4] "We shall place . . . in the exogenous category those variables that are wholly explained by other disciplines" (p. 70).

the measurement of income or wealth, with the meaning of the concepts measured and with the interpretation or analysis of the measures. Most economists and statisticians engaged in studies of income and wealth are concerned with basic problems of measurement, the formulation of concepts and definitions, the accumulation of data from diverse sources ranging from administrative statistics to family living surveys and the adaptation of these data to the conceptual frame-work of a particular system of accounts. The greater number of entries in Volume II, as in Volume I, present national estimates or survey data on income, wealth, and related topics for various dates.

In general, the estimates of national aggregates seem to be conditioned in all countries by the same limitations imposed by the nature of the source material, and the supply of data, and by the financing of the necessary research. The first of these conditions is apparently leading towards some uniformity in the details of the national accounts for various countries under the influence of the various recommendations of the United Nations' Statistical Commission. The second explains the variation in the volume of estimates among countries in possession of the source data and the technical skill required for their utilization in the construction of systems of national accounts.

The discussions of concept in the writings listed in Volume II display, on the other hand, no clear tendencies towards the resolution of the controversies that have filled the literature on national accounts for the past two decades or more. From the viewpoint of a statistician with a bias towards empirical methods the controversial issues can not be settled until they are submitted to a thorough logical analysis. The operational nature of the basic concepts needs clarification; the limitations on measurement should be recognized; and the distinctions between the concept specified for measurement and the theoretical construct of the same name should be sharply drawn. Until the logic of this field of measurement has been considerably refined, the contribution of statisticians to the improvement of the quality of the primary estimates will be seriously restricted. A few entries in Volume II discuss the potential usefulness of sample surveys for the collection of data needed for the purposes of national accounts and the need for developing measures of the error in estimates. The statistician's skills can not, however, be effectively employed until the basic concepts are made to correspond more simply with magnitudes that can be directly observed than is now the case.

Basically the concepts specified in national accounts are synonymous with the operations for determining the particular measurements. National income analysts have adopted the tools of the accountant and are developing rules for their application to data from diverse sources relating to the economic activities of the productive units in the particular country. The productive units fall into three classes, business enterprises, households, and governments. The economic activities of the three classes of units are not distinct

and they overlap in different ways and in varying degrees among the nations now attempting to accumulate national income estimates and other social accounts. The procedures of accounting apply directly to one class of units, business enterprises and such concepts as wealth, income, production, consumption, and capital formation can be differentiated in the records of the economic activities of these units. More or less arbitrary but fairly standardized rules have been devised for fitting such activities of other classes, households and governments, into the same structure of measurements as overlap with or substitute for the activities of business enterprises. Thus the food the farm family consumes that otherwise might be sold by the farm business; the home occupied that might otherwise be rented by the owner as a landlord; the goods and services provided by a public authority that elsewhere are produced by business enterprises can be described by essentially the same structure of accounting concepts.

Difficulties arise when operational definitions are extended well beyond the domain of experience with their application. The Bibliographies indicate clearly that the scholars engaged in the measurement of national accounts were, as of 1949, still unwilling to cope with the fundamental limitations on the extrapolation of concepts beyond the region of their applicability. The blurring of concepts outside of the original area of definition which occurs in all fields of measurement is particularly well illustrated by the problems encountered in describing the economic activities of "subsistence" units for the purposes of national accounts. The differentiation between production and consumption, saving, wealth, income, and similar concepts cannot be simply established by adopting the operational definitions that serve for the aggregation of the accounts of business enterprises. Volume II of the Bibliography, like Volume I, lists very few references to any attempts at measurement of the activities of subsistence units. The small number of entries does not indicate, however, that the analysts have rejected this domain as not susceptible of measurement in terms of prevailing concepts. There are as many references which list the failure to measure production for home consumption as an element affecting the accuracy of comparisons of national income estimates among countries and between dates. The absence of "data," not a consideration of the validity of concepts in this domain, probably explains the small amount of research on this subject.

The case is quite the contrary with government purchases which, beyond the area of activities overlapping with business enterprises, present the difficulties of extrapolation to another scale. Government accounts and balance sheets are not easily summarized to yield the concepts specified for aggregated accounts of a whole economy. There are a fairly large number of direct references, some twenty papers dealing explicitly with the introduction of government activities into national accounts and nearly half of the general papers on concept treat the problem. There are doubtless many ways of explaining the apparent concentration of analysts in the years 1948–49 on the problems involved in fitting government activities into the structure of

national accounts and answering Kuznets' questions in the preface to the Bibliography, Volume I, "Why do 'official' estimators in industrially developed countries in recent years adhere so closely to the treatment of all government purchases as final product, whereas unofficial estimators do not easily accept the underlying assumption? Does the former treatment permit greater ease in planning and budgeting governmental activities and tracing their impact on the short-term economic situation of the country?"

Data in the form of government accounts do exist in most countries and it is not surprising to find the empirical workers concentrating on the problems of extension and adjustment of concepts to this sector of economic activities. The answers to Kuznets' two questions will be given by experience with statistical analysis of the national estimates that have been accumulating in volume sufficient to warrant this Bibliography to cover the output of two years.

The estimates of national product, national income, and other forms of social accounts have become exceedingly important in national policy, planning and forecasting. The tools of the statistician may not, for some time to come, be of much use in the primary processes of data collation and estimation but in the processes of analysis the statistician has much to offer. The real refinements of the concepts in national accounts can be expected to come as experience with statistical analysis of the existing estimates, however incomplete, inaccurate, and inconsistent they may be, reveals the structure of definitions that is practically most useful for the purposes the estimates are intended to serve.

The relative virtues and values of the sundry concepts for prognosis, forecasting and general economic studies will probably be decided by the results of statistical analysis long before the scholars can agree on the identification of the concepts specifying the measurements with theoretical abstractions. The accumulation of observations and methods of empirical analysis may soon render some of the controversial issues obsolete, even meaningless. The stress, for example, on national income or some other aggregate as an absolute measure of "welfare," a really vague abstraction, in the formulation of concepts will doubtless disappear as experience with analysis is extended in time and over the nations of the world. Analysis to date, it is true, can produce little that merits such optimism about the progress of empirical knowledge of the numerous economic magnitudes that fill in the national accounts. In particular, the much debated "consumption-saving function," which accounts for at least 10 papers listed in the Bibliography, Volume II, does not seem to have produced any promising new tools of analysis. Nevertheless, the increasing volume of estimates of national income, its variants and components, national wealth and social accounts has already stimulated new work in time series analysis and surely will lead to some new developments in the utilization in economic analysis of observations on geographic areas. The estimates recorded in the Bibliography for 1948–49 cover more than 70 countries or colonies for varying dates or ranges of dates. Surely this

growing wealth of information, conveniently located through the Bibliographies, will inspire more than a few statistical analysts to surmount the difficulties imposed by variability in concept and in the precision of estimates and devise the means of analysis. Empirical science thrives on the continuous revision of theoretical concepts imposed by experimental findings and the steady refinement of methods of measurement and analysis that result from these revisions.

The Role of Federal Credit Aids in Residential Construction. *Leo Grebler.* New York: National Bureau of Economic Research, Inc., 1953. Occasional Paper No. 39. Pp. 76. $1.00. Paper.

SHERMAN J. MAISEL, *University of California (Berkeley)*

THIS paper is a concise review of the direct impact of the Federal government on residential financing. Data have been gathered from a wide variey of sources. They have been well organized so as to give to the reader a clear picture of the extent of Federal intervention in the real estate financing field. By judicious selection and analysis, the author clearly illuminates the possible results of these governmental policies. While the analysis is clear, lack of data has made it impossible to appraise fully most of the hypotheses concerning effects that are presented. This lack of data, together with the large amount of piecing together of disparate statistics required for this study, is a clear indication of the failure of the Federal housing agencies to collect and publish adequate statistics required for an accurate evaluation of their programs and policies.

The paper points out that under the programs of the Federal Housing Administration and Veterans Administration, the government in recent years has insured or guaranteed between 40 and 50 per cent of loans made on new housing. As a result of these governmental policies, the level of house construction was probably somewhat higher than it would otherwise have been; the market for houses was almost certainly widened; and the costs of ownership may have been lowered. Although primarily as by-product results of actions taken in response to other needs, there have been important impacts on the type of houses built and on institutional practices in the lending markets.

Many readers will find that the value of this review of policy is increased by a supplemental statement of the author as to the conclusions concerning the future that he draws from the past trends. By adding such a statement, the author obviously opens himself to attack from those who disagree with his reading of history. This reviewer feels, however, that the clear statement of future policy problems which may arise from the action taken to date greatly enhances the value of this paper. Providing, as in this case, sufficient detail is given to make disagreement possible, a consideration of future problems by an author completing a detailed analysis of the past is a valuable addition to the study, whether or not one agrees with the conclusions.

Retail Prices and the Consumer Preference. Studies in Business and Economics, Volume 6, Number 1. University of Maryland, College Park, Md.: Bureau of Business and Economic Research, 1952. Pp. 8.

SOPHIA GOGEK, *Standard Oil Company (N. J.)*

THIS pamphlet contains three brief papers: "Price Variations in Men's Clothing by Kind and Location of Store," "Retail Prices, Income and Race," and "Consumer Patronage of Independents and Chains in 59 Cities," each consisting primarily of a summary of the data analyzed. The first compares the prices at neighborhood and centrally located specialty and department stores within and among eighteen cities of eleven items of men's clothing. The second investigates the relative prices in one negro and three white (low, medium, and high rental) neighborhoods in Chicago of a selection of dry goods, men's clothing, and women's clothing. The third summarizes consumer preferences, in shopping for food and for drugs, as between independent and chain stores, and as between neighborhood and centrally located outlets.

Analyses of this type would be both more interesting and more useful if there were more discussion of statistical problems and of the economic and cultural factors responsible for the behavior of the data as recorded. (The second paper, which does go into these matters to some extent, is much the best of the three.) A listing of the sources of data, where they are apparently secondary, would also be helpful.

Industrial Specifications. *E. H. MacNiece.* New York: John Wiley and Sons, 1953. Pp. xiii, 158. $4.50.

J. H. CURTISS, *New York University*

THIS little book gives a description of several different types of written industrial specifications now in use. The classes discussed are raw-material specifications, process specifications, product specifications, and purchase specifications promulgated by public agencies. The exposition is illustrated rather extensively by examples.

The treatment is generally descriptive rather than analytical, and accepts current practices without criticism, constructive or otherwise. The suggestions as to the techniques of preparing specifications are largely confined to listing the topics which should be covered and to urging that clear language be employed.

This serves a useful purpose, and the book should be an influence toward better specifications. On the other hand, the deeper problems of specification-writing are dealt with only briefly and qualitatively. The problems which the reviewer has in mind are exemplified by these: Should acceptance sampling requirements be included in the written specification along with design requirements or should they be issued separately? Should design requirements ever be stated in terms of the behavior of random samples? (In several of MacNiece's examples, this does occur.) How should *laboratory research* data be

processed so as to derive valid tolerance limits? Are inspection data better than laboratory data for this purpose? And so forth. MacNiece handles problems of this sort largely by generalities, such as "Quality control and industrial specifications are blood brothers, and each depends on the other."

But the choice of subject matter is always an author's prerogative. The fact that this little volume is non-technical makes it easy to read, and MacNiece writes fluently and well. Quite apart from its intended use, the book should provide some useful case-history material for teachers of quality control and industrial management.

Analysis in Dental Research. *Neal W. Chilton.* Washington, D. C.: Office Technical Services, U. S. Department of Commerce, 1953. Pp. 216.

GEOFFREY BEALL, *University of Connecticut*

THIS introductory text on statistics is devoted entirely to data obtained in dental studies. One's first impression is to regret that not only must we accept the modern departmentalization of knowledge but must suppose modern man incapable of transferring mental operations from one field of fact to another. On second thought, the dental student, research worker, and clinician will indeed enjoy a book where statistics is so richly illustrated in their terms. The book bears the marks of having been tried and proven in class; the discussion is pleasant and reasonable.

The general treatment shows the mark of being unduly influenced by the extensive and varied literature of texts introducing statistics. Thus, in connection with the calculation of a standard deviation from a sample, the author follows the very old and curious procedure of recommending a biased estimate, having divisor N, when N is greater than 30. On the other hand, as in connection with concepts of chance variation and the binomial distribution he is influenced by the vogue of the moment which would circumscribe statistics to an elaboration of combinatorial probability. This confusion of position does not help general organization. Thus correlation is presented in an infelicitous manner involving the estimation of standard deviations and means. It should, in the present context, be reduced to a form involving only sums of various kinds. An unhappy mixture occurs in the discussion on frequency graphs together with the general nature of numbers and the presentation of time series.

Various particular faults may be found. Thus the discussion of the mode is confusing since the term is not used in the ordinary statistical sense but here means the most common class. The author gives a lengthy and involved discussion of "normal curve analysis of the four-fold table" without apparently realizing that he has calculated a value of chi (not squared). He then goes on to presentation of the contingency for the 2×2 table to get chi-squared. In neither case does he include Yates' correction for discontinuity.

In spite of faults the book still has practical virtues. It concerns itself first with the analysis of qualitative data and then with quantitative data as a less

common even if more happy situation. The author does not stress the *t*-distribution but considers tests of various kinds on a simple normal basis, although he devotes a terminal chapter to a very satisfactory discussion of the *t*-test and analysis of variance.

Intrastate Migration in Michigan: 1935–1940. *Amos H. Hawley.* Michigan Governmental Studies No. 25: Ann Arbor: University of Michigan Press, 1953. Pp. 199. $1.50. Paper.

JOHN FOLGER, *Southern Regional Education Board*

THIS is a book of facts about migrants in Michigan. As an example of pure empiricism it is outstanding, description hardly ever giving way to analysis or interpretation. A good index of the descriptive nature of the work is the fact that it contains 106 tables, 28 figures, and only one reference to any other study in the field of migration.

Hawley is to be commended for the workmanlike job of reducing the tremendous amount of detailed data contained in the subregional tabulations of migration to understandable proportions. But the study never goes beyond the initial framework provided in the Census data. No interpretation is given either by comparison with migration in other times or places or in terms of conditions or changes in the subregion of origin or destination. The student of population seeking new theories of migration or new tests of old hypotheses will be disappointed. The facts are there, and all of the pleasure and hard work of assessing their meaning is left to the reader.

The measures of migration used are all easily understood, proportions and simple rates being employed throughout most of the book. Greater use of more refined rates and indirect standardization techniques would have been valuable in removing the effects of differences in characteristics between migrants and nonmigrants when studying the differences in other characteristics. For example, how much of the difference between migrants and nonmigrants in marital status is attributable to age differences between the migrants and the total population? While adjusted rates and proportions are harder for the reader to comprehend, they are almost a necessity in the process of untangling the complex of variables associated with the process of migration.

It is this reviewer's hope that future studies of migration utilizing the rich detail of the subregional tabulations will be able to perform both the competent summarization of the data that Hawley has provided, and a systematic testing of hypotheses about migration and the characteristics of migrants.

Sample Surveys of Current Interest. (Fourth Report.) Statistical Papers, Series C, No. 5. New York: Department of Economic Affairs, Statistical Office of the United Nations, March 1952. Pp. 56. Paper. 50 cents.

HAROLD NISSELSON, *Bureau of the Census*

THIS Report is the fourth and latest in a more or less annual series begun in 1948 at the suggestion of the Sub-Commission on Statistical Sampling

of the United Nations Statistical Commission, "to assist those interested in the application of modern sampling techniques by making available, in summary form, the experiences gained by various statisticians and statistical organizations in the field." Altogether, 33 of the 60 U.N. countries are represented in one way or another in the four Reports published to date. Perhaps not unexpectedly, there have been no reports from Russia nor from any of the "Iron Curtain" countries except Czechoslovakia (twice), Hungary (twice) and Poland (once). The projects reported are presumably not to be taken as a proper sample of all sample surveys being carried out in the 60 U.N. countries. However, the general impression left by the Reports is that the ideas and techniques of modern probability sampling have achieved a rapid and wide circulation, both geographically and in subject matter dealt with. On the other hand, hardly any reference is made to non-sampling errors in surveys. It is clear that elsewhere, as in the U.S.A., the general theory of survey design—in contrast to its sampling phases—is in a relatively undeveloped state.

The information for these Reports is obtained by circularizing the national statistical offices of all U.N. countries (in the U.S., the Division of Statistical Standards, Bureau of the Budget, Executive Office of the President), with dependence on the cooperation of those offices for replies. The reports, therefore, have three characteristics. First, they appear to include only work of official and quasi-official (e.g., The Indian Statistical Institute) organizations. Second, not all current projects are reported, as is explicitly noted for India, the U.K., and the U.S.A., and the selections may be considered somewhat arbitrary. Third, the project descriptions do not completely adhere to the recommendations of the Sub-Commission itself.[1] Information such as variances of sampling units or unit costs is scanty; and the summaries tend to be long on description and short on evaluation. (The U.S.A. is not outstanding in these respects.) Statisticians interested in technical details of the design of a given project, or in a realistic evaluation of the experience in carrying the project out, will therefore generally need to write to the United Nations Statistical Office or the sponsoring organization for further information. Nevertheless, the Reports do provide a compilation in a single place of some of the more important official sampling projects being carried on in various countries, many of which have not been published as papers on sampling. For example, the present Report contains the only published statement currently available on the applications of sampling in the enumeration and processing of the 1950 Censuses of Population and Housing. For this purpose the Reports are more comprehensive, and should be more useful, than specialized listings presently appearing in scattered journals. It presumably lies with statisticians themselves to decide how useful this series may be; and, by their cooperation to remedy any short-comings in the Reports.

[1] "The Preparation of Sampling Survey Reports," Statistical Papers, Series C, No. 1, Statistical Office of the United Nations.

Age and Achievement. *Harvey C. Lehman.* Princeton, N. J.: Princeton University Press, 1953. Pp. xiii, 359. $7.50.

LEONA E. TYLER, *University of Oregon*

Too often in psychology research workers have made generalizations that extend far beyond the evidence their data present. This book is a striking exception. Its 331 pages with their 170 graphs and 61 tables are all related in some way to the important question, "What are man's most creative years?"

What the author has done is to analyze facts reported in standard reference books with regard to the ages at which great men have made their outstanding creative contributions. Science, philosophy, music, art, and literature have all come under his scrutiny. The result is a large number of curves, each of which shows the average number of important contributions for each five-year age period in the area of human achievement to which it applies. All of these curves are placed on the same scale by the simple device of reducing the figures to percentages. Whatever its absolute value, the highest number, representing the peak age interval, becomes 100 per cent and the others take lesser values accordingly. Other statistical information is given in extensive tables so that the reader can quite easily locate numbers of cases, medians, means, and standard deviations for the various distributions. The discussion emphasizes the modal values, the peaks of the curves, more than anything else, but the other figures are not ignored. The derivation of basic data from biographical dictionaries insures freedom from bias since such figures were compiled by men who were not concerned at all with this particular question.

In addition to his discussions of the central problem, the author includes some chapters on supplementary issues, such as the relationship between early achievement and total output, and comparisons of age trends in earlier and later periods of history.

Several conclusions are drawn from the data presented. First of all, the vast majority of the curves show a remarkable similarity. The peak occurs early, most commonly in the 30's, but as low as the 26–30 interval for chemistry and as high as the 40–44 interval for light opera and musical comedy, for "best books," and for metaphysics. All curves show a gradual decline, with some achievement indicated even into advanced ages. It is perhaps noteworthy that the data for sports and for such activities as chess show the same trends as the others. Championships are won most frequently by persons under 36.

Secondly, when the works agreed by most critics to be of the very *highest* quality are singled out for special attention, the curves drawn from these figures show earlier peaks and a more rapid descent than do those for the less selective distributions. This finding is stressed again and again. The more *quality* as distinguished from quantity is considered, the younger is the average age at which it is achieved.

Thirdly, data for incomes and for positions of leadership produce curves

which are quite different from those for creative achievement. Their peaks occur most commonly at ages of 50 and higher.

The author does not present any one type of explanation for these findings. He reminds us periodically that these are not biological curves representing the rise and decline of vigor or intelligence in any *one* individual, although if there were such developmental trends they would reflect themselves in this way. The extensive documentation in Chapters 13 and 14 with regard to outstanding contributions made at very early and at very advanced ages, as well as the evidence that trends have changed somewhat from one historical period to another, argue against a *purely* biological explanation. There are a number of other possible causes centering around motivation and around the circumstances in which great men typically find themselves at different periods of life. These are listed in detail in the last chapter.

It is possible to criticize Lehman's report on two counts, one rather superficial, the other less so. First, it would have made better reading if the author had worked over and reorganized his material instead of simply combining into a book chapters that had already been published as journal papers. There is a great deal of repetition from chapter to chapter, as things stand, and the same points are made again and again. Second, the statistical work is confined entirely to *descriptive* statistics. Our judgments must be based on a scrutiny of the figures and tables, without benefit of significance tests of any kind. So far as the main question is concerned, this is probably satisfactory, since the position of the "peaks" is easily determined by inspection. But other questions would seem to call for tests of the statistical significance of differences. How similar are various curves in shape? Is there really a tendency toward bimodality as many of the figures in the first few chapters would suggest? Is the change from an earlier to a later period, as discussed in Chapter 18, significant? It is unfortunate that Lehman's interest in the *shape* of the age curve has in some instances led him to give us data to which significance tests cannot very well be applied. For instance, in order to show peaks clearly, five-year intervals were selected in different ways for different bodies of data. Thus for German lyrics and ballads, the intervals are 17–21, 22–26, 27–31, etc. instead of the 20–24, 25–29, 30–34, etc. which are used in most of the tabulations. Distributions set up in this way with intervals chosen after inspection would not be statistically comparable. Another instance occurs in Chapter 18 where the conclusions about historical periods would be strengthened by significance tests. What Lehman has done in each of his groups is to compare curves for the 50 per cent born earlier and the 50 per cent born later. The result is that the actual historical period represented varies from one comparison to the next. For example, a geologist born in 1800 is included in the "early" group, whereas a mathematician born in the same year appears in the "later" group. Thus different groups of scientists could not very well be combined.

These criticisms, it is true, apply to only a fraction of the data presented. Other workers wishing to test various hypotheses about age trends in achieve-

ment will find this book a mine of information. Much use can probably be made of means and standard deviations as given in the various tables. While we can wish that the author had gone further in using this large body of data to enlighten us as to *why* these trends appear, we can at the same time be grateful to him for digging out the figures which show the trends themselves so plainly.

Backgrounds of Human Fertility in Puerto Rico: A Sociological Survey. *Paul K. Hatt.* Princeton: Princeton University Press, 1952. Pp. xxiv, 512. $5.00. Paper.

HENRY S. SHRYOCK, JR., *Bureau of the Census*

THIS book is the first and most general report of a field survey conducted in 1947 under the auspices of the Office of Population Research, Princeton University, and the Social Science Research Center, The University of Puerto Rico. The study was designed to "throw light upon those basic attitude patterns and life conditions which affect fertility levels in Puerto Rico." Its conceptual design has much in common with the 1941 Indianapolis Study of Social and Psychological Factors Affecting Fertility.

The probability sample used by the Insular Bureau of Labor Statistics in its monthly survey of the labor force provided a prelisting from which the present sample was drawn. Of all listed households, 92 per cent were located and at least one adult interviewed. Within these households 92 per cent of the adults were interviewed so that about 85 per cent of the adults covered by the original list were probably interviewed. The actual sample was checked for representativeness against the 1940 Census and by comparing households missed with neighboring households. There were moderate biases in the direction of the exclusion of youths and the economic extremes. (Data from the 1950 Census, which have subsequently become available, indicate somewhat less bias in the sample with respect to the sex and age of adults.)

The supervisors of the field work were social science graduates. The 65 experienced interviewers, all of them women, were given two days' training plus a conference after two practice interviews. The schedule was pretested upon Puerto Rican women living in New York. The final forms, as well as the instructions to the enumerators, are reproduced in appendices, although only in the original Spanish. This rather long and intensive schedule consisted of several parts: form 1 applying to each household, form 2 to each person 15 years old and over, and forms 3 and 4 to each person who had had any kind of marital experience (legal, consensual, or concubinal). On the other hand, the 1950 Census question on children ever born was asked of all women, including the single; and the over-all nonresponse rate was quite low.

After an initial chapter on "Methodology," the next four chapters are entitled, "The General Patterns of Social Conditions and Social Attitudes," "Socio-economic Status: Rental Value and Educational Level," "Rural-

urban: Birth and Residence," and "Age Differences," respectively. These chapters treat the relationships among various background factors but do not deal with measures of actual fertility. The relationships are examined by tetrachoric correlation, with partial and multiple as well as zero order coefficients being presented. In many instances there would seem to have been enough cases to permit further cross-classification, which would have been more satisfactory than holding a factor "constant" through partial correlation.

The factors brought into this correlation analysis include attitudes toward a number of social questions, including, among others, the desirability of consensual marriage for both men and women, ideal age at marriage, and ideal size of completed family. The attitude measures themselves are examined in several interesting ways (for example, ideal age at marriage is compared with own actual age at marriage and number of children desired for daughter is compared with the woman's own actual fertility). The attitude questions and the construction of the measures based on them seem relatively primitive after the reports of the Research Branch, Information and Education Division, War Department, although it must be allowed that little had been published by Stouffer and his associates on their methodology by the time the present study was designed. (The reviewer suspects that the rather disappointing findings of the Indianapolis Study concerning the effects of psychological factors on fertility would be replaced by much more positive results could this new apparatus of attitude research be employed.)

Tetrachoric correlation is also employed extensively in the chapter on "Differential Fertility," which contains much of the meat of the study. It seems unfortunate that, for mechanical considerations of processing the data, fertility is expressed in terms of pregnancies rather than of live births, since the latter were probably reported more accurately. Robert Osborn, Jr., has contributed a long chapter on "The Trend in Fertility," using mostly cohort analysis of the survey data but also pertinent census data and vital statistics. The downward trend that he is trying to measure seems to be a recent and slight one—less pronounced, for example, than the downward trend of fertility in the British Caribbean possessions. This kind of job needs to be done, but it is often thankless work because many readers are frustrated by the absence of definite conclusions as to the size of the various suspected biases and their net effect. Osborn is dealing with data that just do not yield such conclusive answers, however.

In conclusion, this book represents a sound job by the late Professor Hatt and his associates. Their findings are interesting and significant. Insular attitudes toward such things as family limitation, ideal size of family, ideal age at marriage, education wanted for children, and the employment of women outside the home are surprisingly like those on the mainland but also greatly at variance with current practices in Puerto Rico. If only a part of these changing attitudes can be translated by the oncoming generation into

actual conduct, there would be a very substantial drop in the high fertility levels now prevalent. From the standpoint of statistical methodology, this study contains few innovations; but the report does demonstrate that a fairly complicated kind of sample survey can be successfully carried out in this kind of underdeveloped area.

An Approach to Measuring Results in Social Work. *David G. French.* New York: Columbia University Press, 1952. Pp. xiv, 178. $3.00.

JOHN E. WALSH, *U. S. Naval Ordnance Test Station*

THIS book is a report on the 1950 Michigan Reconnaissance Study of Evaluative Research in Social Work, which was sponsored by the Michigan Welfare League and financially supported by the James Foster Foundation. "Evaluative research" refers to research performed to evaluate a practice or a policy. The book is not itself a piece of evaluative research but rather a survey of the possibilities in applying evaluative research to social work. Although carried out with reference to social work in Michigan, the analysis and recommendations are believed to have general applicability. The presentation is both logical and lucid.

Three general procedures were used in performing the Reconnaissance Study. First, factual information which appeared to have application to the planning of a research program in social work was assembled for the field of research and for the field of social work. Second, the available literature dealing with the general field of applied research, with emphasis on social science research and social work, was reviewed. Finally, both individual and group conferences were held to supplement the available printed information. These conferences were used to obtain suggestions and observations as well as to test portions of the analysis and conclusions developed during the study. As the study progressed, the emphasis shifted from evaluation to research. This was due to the great difficulties involved in performing valid and useful research in the social work field. However, the evaluation aspect was retained as a meeting place for the practical concerns of the social worker and the more theoretical attitude of the researcher.

An outline of the content of the seven chapters of the book furnishes a good indication of the approach used in the Reconnaissance Study. Chapter I describes the general motivation which led to the study. Chapter II considers the situation in Michigan and emphasizes the large financial investment in social work (approximately $125,000,000 in Michigan for 1950). Chapter III states some general types of questions which professional and lay persons raise concerning social work and studies these questions from the viewpoint of planning a research program. Chapter IV analyzes what is involved in doing evaluative research. Four representative evaluation studies in social work were reviewed to determine their strong and weak points. Chapter V considers the relationship between theoretical and applied re-

search as a basis for planning a research program in social work. Chapter VI states eight criteria which are believed to be important for any continuing program of research in social work and outlines an administrative setting which appears capable of satisfying all the criteria. Chapter VII extends the general considerations of the preceding chapter into some moderately specific suggestions concerning the establishment of an institute for research in social work.

The four methodology analyses of evaluation studies considered in Chapter IV are themselves of interest. Only the study title, the authors, and the person making the analysis are stated here: (1) "Measuring Results in Social Casework: A Manual on Judging Movement," by J. McVicker Hunt and Leonard S. Kogan with analysis by John G. Hill. (2) "Changing Attitudes Through Social Contact," by Leon Festinger and Harold Kelley with analysis by Leon Festinger. (3) "An Experiment in the Prevention of Delinquency," by Edwin Powers and Helen L. Witmer with analysis by Helen L. Witmer. (4) "Unraveling Juvenile Delinquency," by Sheldon and Eleanor Glueck with analysis by Alfred J. Kahn.

The principal recommendation of the book is that an appropriate location for research in social work is as an institute which is part of a graduate professional school of social work in a university. This recommendation is motivated by the criteria of Chapter VI and seems to be both reasonable and feasible. The specific suggestions and financial estimates given in Chapter VII for the establishment and staffing of such an institute for research in social work also appear to be reasonable.

This book represents a well-thought-out first approach to the problem of obtaining valid quantitative information in the important practical field of social work. The basic difficulties are clearly recognized and no attempt is made to underemphasize their magnitude. Slow progress is anticipated in attaining even moderate objectives.

Evidently the development of research practices and theories which have practical value for social work would require a large amount of time, effort, and money. However, the great scope and importance of social work would seem to indicate that the eventual gains from a well-founded evaluative research program would far offset the investment involved. This has been the case in many other practical fields and there is no indication that the social work field represents an exception.

RANDOM DIGITS (15,126–17,375)

From *A Million Random Digits*, to be published by the Rand Corporation, Santa Monica, California.

33591	59785	12833	98932	68064
58418	90331	55858	04015	21454
64446	51017	22280	75597	50227
72136	00303	38880	93327	49522
98626	82484	54610	07211	78610
58393	20225	05436	46172	88951
37346	51007	38032	36002	21080
70712	44236	96795	92351	92844
93585	09918	30983	44282	66849
09473	72923	16747	49802	50639
40229	34921	60405	06803	19332
39795	77221	10012	40798	33864
10288	57483	10881	58984	45136
75702	69428	34047	76224	45887
18129	93659	58389	19715	66259
17777	41004	47057	30688	07539
75195	62294	03371	11672	13089
09722	67635	12114	63055	09214
78800	86912	42076	50287	97998
94287	54751	36242	36557	85664
35997	30761	97081	09501	68887
88241	30402	12318	52430	40139
54382	73370	26184	14024	57444
77681	74946	02099	69091	19372
53148	26074	52293	65359	63971
27212	89889	46933	13364	33883
03867	03105	87912	29610	75108
79895	82633	19209	21548	35022
55256	69386	57453	70147	73538
75937	31113	07607	48037	71020
83389	80236	65972	74528	40888
37363	30345	79933	71058	34826
21960	95585	40374	13239	56162
49562	44137	46625	20031	08524
22666	27414	30980	74485	26480
55386	95918	92481	49234	62616
97725	69513	36950	63526	93824
04703	95851	90956	64424	95979
20077	65817	99523	73180	59978
21343	30031	18840	99260	21284

54021	29008	83672	53679	96395
71820	11033	20183	08804	29493
17709	94849	31771	23244	81585
08254	78963	95437	30231	59108
81351	70961	10255	60437	84576
30665	43699	03593	29165	59990
09069	78653	90094	42735	82876
50593	14698	04737	72551	36417
49578	18100	59836	73221	91696
58388	40915	94507	32209	11548
36995	36202	07971	67001	62062
66348	87666	78055	44485	82955
45457	78252	98239	40000	75563
35406	59553	57852	07506	00009
09678	24538	52426	84852	83781
31604	55850	25644	44972	62275
05523	34355	75127	69797	71419
09458	29207	43632	32905	38513
74086	24077	21369	93541	75329
89299	30765	00348	01134	71581
78089	36887	65867	08595	47390
62067	41248	78667	95282	05622
18924	44409	45345	04972	52794
78656	18550	37171	13483	56058
68420	16483	26482	85964	71336
39182	51174	98146	42953	06606
26224	19972	13442	69662	76755
15737	00496	57764	52052	12835
51093	62290	38948	47473	16618
67799	28342	14551	44149	29854
23875	56766	01932	90063	31883
13701	95168	13169	42055	75609
37741	86434	22400	92732	99851
42479	47405	14055	75427	34464
76691	33263	62048	70917	97697
86707	18895	81790	71294	74832
92564	39987	02283	89970	28776
94648	05598	32171	28793	36746
29838	10664	28050	60122	13409
36368	17792	84792	76594	74110
09254	07510	51039	91683	84500
41486	58524	54508	20707	58504
90794	51934	03295	26582	16300
58558	84833	17105	46659	25003
52392	53546	70291	98846	67315
46095	19501	56181	85351	05023
09925	96974	34321	05454	12862
70550	42483	71204	99628	40642
50281	83708	27298	92651	95086
22767	26514	56322	78635	63731

Princeton UNIVERSITY PRESS

COHORT FERTILITY • Native White Women in the United States •

by P. K. Whelpton

The annual fertility experience from 1915 to 1950 is tabulated and analyzed for the white women born in each of the years from 1880 to 1933. This book develops new techniques of measuring past fertility trends and provides an improved basis for forecasting future trends. *Sponsored by the Scripps Foundation.*

516 PAGES. $6

A THEORY OF ECONOMIC-DEMOGRAPHIC DEVELOPMENT • *by Harvey Leibenstein, foreword by Frank Notestein*

How can the overpopulated, underdeveloped areas of the world break out of the Malthusian trap? This is the first book to deal with the economic aspects of this vital question. No other book yet published attempts to integrate capital accumulation and population growth into a single scheme.

222 PAGES. $4

COLONIAL DEVELOPMENT AND POPULATION IN TAIWAN • *by George W. Barclay, foreword by Frank Notestein*

Offers an unusual and authoritative view of an agrarian region in the process of development by a colonial power. Based on the remarkable and unprecedented statistical data which the Japanese compiled to aid their administration—one of the most complete and creditable records for a population of this size that has ever been at the disposal of demographers.

296 PAGES. 78 TABLES, 32 FIGURES. $5

Order from your bookstore • PRINCETON UNIVERSITY PRESS

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*